(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 December 2007 (13.12.2007)

PCT

(10) International Publication Number
WO 2007/142621 A1

(51) International Patent Classification:
G06K 9/00 (2006.01) *G06K 9/32* (2006.01)
G06K 9/34 (2006.01) *G06K 9/40* (2006.01)
G06K 9/36 (2006.01)

(21) International Application Number:
PCT/US2006/021393

(22) International Filing Date: 2 June 2006 (02.06.2006)

(25) Filing Language: English

(26) Publication Language: English

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 10/608,784 (CIP)
Filed on 26 June 2003 (26.06.2003)

(71) Applicant *(for all designated States except US)*: **FOTONATION VISION LIMITED** [IE/IE]; Galway Businges Park, Dangan, Galway City, County Galway (IE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **STEINBERG, Eran** [US/US]; 137 Granville Way, San Francisco, CA 94127 (US). **CORCORAN, Peter** [IE/IE]; 137 Granville Way, Cregg, Claregalway, Co. Galway (IE). **PRILUTSKY, Yury** [US/US]; 1626 Shoal Drive, San Mateo, CA 94404 (US). **BIGIOI, Petronel** [IE/IE]; 57 Sceilg Ard, Headford Road, Galway, Co. Galway (IE).

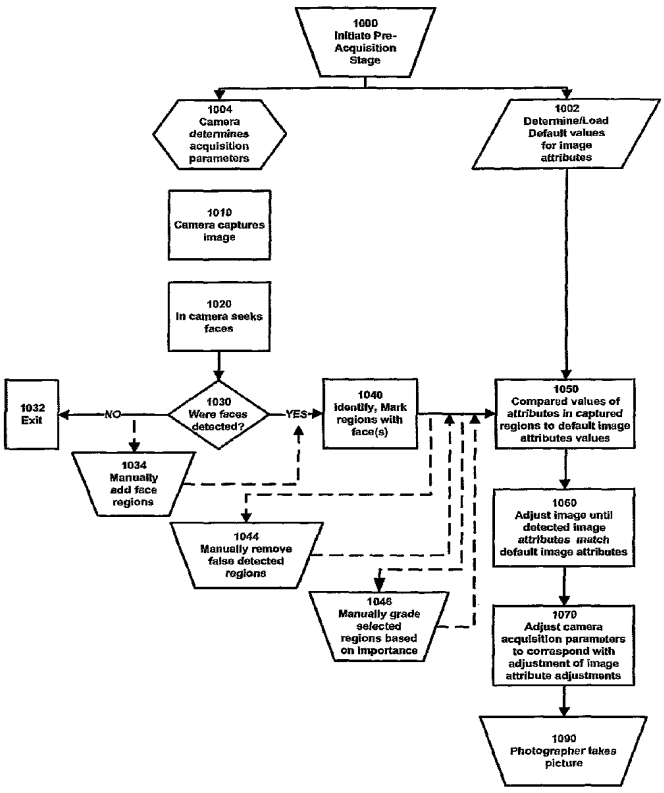
(74) Agent: **SMITH, Andrew, V.**; Jackson & Co., LLP, 6114 La Salle Ave., #507, Oakland, CA 94611-2802 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US (patent), UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH,

[Continued on next page]

(54) Title: MODIFICATION OF POST-VIEWING PARAMETERS FOR DIGITAL IMAGES USING IMAGE REGION OR FEATURE INFORMATION



(57) Abstract: A method of modifying the viewing parameters of digital images using image feature detection for achieving desired parameters based on one or more sub-groups of pixels that correspond to one or more selected image features such as a face, body, pets, or a background region or a foreground region. Such methods may be used for animating still images, automating and streamlining application such as the creation of slide shows and screen savers of images containing selected image features. Two or more features such as a foreground region and a background region may be separated and independently processed.

WO 2007/142621 A1

GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

MODIFICATION OF POST-VIEWING PARAMETERS FOR DIGITAL IMAGES USING IMAGE REGION OR FEATURE INFORMATION

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a CIP of United States patent application no. 10/608,784, filed June 26, 2003, which is one of a series of contemporaneously-filed patent applications including Atty docket 2100874-991210 (FN102-A) entitled, "Digital Image Processing Using Face Detection Information", by inventors Eran Steinberg, Yuri Prilutsky, Peter Corcoran, and Petronel Bigioi; Atty docket 2100874-991220 (FN102-B) entitled, "Perfecting of Digital Image Capture Parameters Within Acquisition Devices Using Face Detection", by inventors Eran Steinberg, Yuri Prilutsky, Peter Corcoran, and Petronel Bigioi; Atty docket 2100874-991230 (FN102-C) entitled, "Perfecting the Optics Within a Digital Image Acquisition Device Using Face Detection", by inventors Eran Steinberg, Yuri Prilutsky, Peter Corcoran, and Petronel Bigioi; Atty docket 2100874-991240 (FN102-D) entitled, "Perfecting the Effect of Flash Within an Image Acquisition Device Using Face Detection", by inventors Eran Steinberg, Yuri Prilutsky, Peter Corcoran, and Petronel Bigioi; Atty docket 2100874-991250 (FN102-E) entitled, "A Method of Improving Orientation and Color Balance of Digital Images Using Face Detection Information", by inventors Eran Steinberg, Yuri Prilutsky, Peter Corcoran, and Petronel Bigioi; Atty docket 2100874-991260 (FN102-F) entitled, "Modification of Viewing Parameters for Digital Images Using Face Detection Information", by inventors Eran Steinberg, Yuri Prilutsky, Peter Corcoran, and Petronel Bigioi; Atty docket 2100874-991270 (FN102-G) entitled, "Digital Image Processing Composition Using Face Detection Information", by inventor Eran Steinberg; Atty docket 2100874-991280 (FN102-H) entitled, "Digital Image Adjustable Compression and Resolution Using Face Detection Information" by inventors Eran Steinberg, Yuri Prilutsky, Peter Corcoran, and Petronel Bigioi; and Atty docket 2100874-991290 (FN102-I) entitled, "Perfecting

of Digital Image Rendering Parameters Within Rendering Devices Using Face Detection" by inventors Eran Steinberg, Yuri Prilutsky, Peter Corcoran, and Petronel Bigioi.

BACKGROUND

1. Field of the Invention

The invention relates to digital image processing and viewing, particularly automatic suggesting or processing of enhancements of a digital image using information gained from identifying and analyzing regions within an image or features appearing within the image, particularly for creating post acquisition slide shows. The invention provides automated image analysis and processing methods and tools for photographs taken and/or images detected, acquired or captured in digital form or converted to digital form, or rendered from digital form to a soft or hard copy medium by using information about the regions or features in the photographs and/or images.

2. Description of the Related art

This invention relates to finding and defining regions of interest (ROI) in an acquired image. In many cases the interest relates to items in the foreground of an image. In addition, and particularly for consumer photography, the ROI relates to human subjects and in particular, faces.

Although well-known, the problem of face detection has not received a great deal of attention from researchers. Most conventional techniques concentrate on face recognition, assuming that a region of an image containing a single face has already been extracted and will be provided as an input. Such techniques are unable to detect faces against complex backgrounds or when there are multiple occurrences in an image. For all of the image enhancement techniques introduced below and others as may be described herein or understood by those skilled in the art, it is desired to make use of the data obtained from face detection processes for suggesting options for improving digital images or for automatically improving or enhancing quality of digital images.

Yang et al., IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 1, pages 34-58, give a useful and comprehensive review of face detection techniques January 2002. These authors discuss various methods of face detection which may be divided into four main categories: (i) knowledge-based methods; (ii) feature-invariant approaches, including the identification of facial features, texture and skin color; (iii) template matching methods, both fixed and deformable and (iv) appearance based methods, including eigenface techniques, statistical distribution based methods and neural network approaches. They also discuss a number of the main applications for face detections technology. It is recognized in the present invention that none of this prior art describes or suggests using detection and knowledge of faces in images to create and/or use tools for the enhancement or correction of the images.

a. Faces as Subject Matter

Human faces may well be by far the most photographed subject matter for the amateur and professional photographer. In addition, the human visual system is very sensitive to faces in terms of skin tone colors. Also, in experiments performed by tracking the eye movement of the subjects, with an image that includes a human being, subjects tend to focus first and foremost on the face and in particular the eyes, and only later search the image around the figure. By default, when a picture includes a human figure and in particular a face, the face becomes the main object of the image. Thus, many artists and art teachers emphasize the location of the human figure and the face in particular to be an important part of a pleasing composition. For example, some teach to position faces around the "Golden Ratio", also known as the "divine proportion" in the Renaissance period, or PHI, φ-lines. Some famous artists whose work repeatedly depict this composition are Leonardo Da-Vinci, Georges Seurat and Salvador Dali.

In addition, the faces themselves, not just the location of the faces in an image, have similar "divine proportion" characteristics. The head forms a golden rectangle with the eyes at its midpoint; the mouth and nose are each placed at golden sections of distance between the eyes and the bottom on the chin etc. etc.

b. Color and Exposure of Faces

While the human visual system is tolerant to shifts in color balance, the human skin tone is one area where the tolerance is somewhat limited and is accepted primarily only around the luminance axis, which is a main varying factor between skin tones of faces of people of different races or ethnic backgrounds. A knowledge of faces can provide an important advantage in methods of suggesting or automatically correcting an overall color balance of an image, as well as providing pleasing images after correction.

c. Auto Focus

Auto focusing is a popular feature among professional and amateur photographers alike. There are various ways to determine a region of focus. Some cameras use a center-weighted approach, while others allow the user to manually select the region. In most cases, it is the intention of the photographer to focus on the faces photographed, regardless of their location in the image. Other more sophisticated techniques include an attempt to guess the important regions of the image by determining the exact location where the photographer's eye is looking. It is desired to provide advantageous auto focus techniques which can focus on what is considered the important subject in the image

d. Fill-Flash

Another useful feature particularly for the amateur photographer is fill-flash mode. In this mode, objects close to the camera may receive a boost in their exposure using artificial light such as a flash, while far away objects which are not effected by the flash are exposed using available light. It is desired to have an advantageous technique which automatically provides image enhancements or suggested options using fill flash to add light to faces in the foreground which are in the shadow or shot with back light.

e. Orientation

The camera can be held horizontally or vertically when the picture is taken, creating what is referred to as a landscape mode or portrait mode, respectively. When viewing images, it is preferable to determine ahead of time the orientation of the camera at acquisition, thus eliminating a step of rotating the image and automatically orienting the image. The system may try to determine if the image was shot horizontally, which is also referred to as landscape format, where the width is larger than the height of an image, or vertically, also referred to as portrait mode, where the height of the image is larger than the width. Techniques may be used to determine an orientation of an image. Primarily these techniques include either recording the camera orientation at an acquisition time using an in camera mechanical indicator or attempting to analyze image content post-acquisition. In-camera methods, although providing precision, use additional hardware and sometimes movable hardware components which can increase the price of the camera and add a potential maintenance challenge. However, post-acquisition analysis may not generally provide sufficient precision. Knowledge of location, size and orientation of faces in a photograph, a computerized system can offer powerful automatic tools to enhance and correct such images or to provide options for enhancing and correcting images.

f. Color Correction

Automatic color correction can involve adding or removing a color cast to or from an image. Such cast can be created for many reasons including the film or CCD being calibrated to one light source, such as daylight, while the lighting condition at the time of image detection may be different, for example, cool-white fluorescent. In this example, an image can tend to have a greenish cast that it will be desired to be removed. It is desired to have automatically generated or suggested color correction techniques for use with digital image enhancement processing.

g. Cropping

Automatic cropping may be performed on an image to create a more pleasing composition of an image. It is desired to have automatic image processing techniques for generating or suggesting more balanced image compositions using cropping.

h. Rendering

When an image is being rendered for printing or display, it undergoes operation as color conversion, contrast enhancement, cropping and/or resizing to accommodate the physical characteristics of the rendering device. Such characteristic may be a limited color gamut, a restricted aspect ratio, a restricted display orientation, fixed contrast ratio, etc. It is desired to have automatic image processing techniques for improving the rendering of images.

i. Compression and resolution

An image can be locally compressed in accordance with a preferred embodiment herein, so that specific regions may have a higher quality compression which involves a lower compression rate. It is desired to have an advantageous technique for determining and/or selecting regions of importance that may be maintained with low compression or high resolution compared with regions determined and/or selected to have less importance in the image.

SUMMARY OF THE INVENTION

A method of generating one or more new digital images, or generating a progression or sequence of related images in a form of a movie clip, using an original digitally-acquired image including a selected image feature is provided. The method includes identifying one or more groups of pixels that correspond to a selected image feature, or image region within an original digitally-acquired image. A portion of the original image is selected that includes the one or more groups of pixels segmented spatially or by value. Values of pixels of one or more new

images are automatically generated based on the selected portion in a manner which includes the selected image feature within the one or more new images.

The selected image feature may include a segmentation of the image to two portions, e.g., a foreground region and a background region, and the method may include visually separating the foreground region and the background region within the one or more new images. The visual encoding of such separation may be done gradually, thereby creating a movie-like effect.

The method may also include calculating a depth map of the background region. The foreground and background regions may be independently processed. One or more of the new images may include an independently processed background region or foreground region or both. The independent processing may include gradual or continuous change between an original state and a final state using one of or any combination of the following effects: focusing, saturating, pixilating, sharpening, zooming, panning, tilting, geometrically distorting, cropping, exposing or combinations thereof. The method may also include determining a relevance or importance, or both, of the foreground region or the background region, or both.

The method may also include identifying one or more groups of pixels that correspond to two or more selected image features within the original digitally-acquired image. The automatic generating of pixel values may be in a manner which includes at least one of the two or more selected image features within the one or more new images or a panning intermediate image between two of the selected image features, or a combination thereof.

The method may also include automatically providing an option for generating the values of pixels of one or more new images based on the selected portion in a manner which includes the selected image feature within each of the one or more new images.

A method of generating one or more new digital images using an original digitally-acquired image including separating background and foreground regions is provided. The method includes identifying one or more groups of pixels that correspond to a background region or a foreground region, or both, within an original digitally-acquired image. The foreground portion may be based on the identification of well known objects such as faces, human bodies, animals and in particular pets. Alternatively, the foreground portion may be determined based

on a pixel analysis with information such as chroma, overall exposure and local sharpness. Segmentations based on local analysis of the content or the values may be alternatively performed as understood by those skilled in the art of image segmentation. A portion of the original image is selected that includes the one or more groups of pixels. Values of pixels of one or more new images are automatically generated based on the selected portion in a manner which includes the background region or the foreground region, or both. The method may also include calculating a depth map of the background region. The foreground and background regions may be independently processed for generating new images.

The present invention and/or preferred or alternative embodiments thereof can be advantageously combined with features of parent U.S. patent application serial number 10/608,784, including a method of generating one or more new digital images, as well as a continuous sequence of images, using an original digitally-acquired image including a face is further provided. A group of pixels that correspond to a face within the original digitally-acquired image is identified. A portion of the original image is selected to include the group of pixels. Values of pixels of one or more new images based on the selected portion are automatically generated, or an option to generate them is provided, in a manner which always includes the face within the one or more new images.

A transformation may be gradually displayed between the original digitally-acquired image and one or more new images. Parameters of said transformation may be adjusted between the original digitally-acquired image and one or more new images. Parameters of the transformation between the original digitally-acquired image and one or more new images may be selected from a set of at least one or more criteria including timing or blending or a combination thereof. The blending may vary between the various segmented regions of an image, and can include dissolving, flying, swirling, appearing, flashing, or screening, or combinations thereof.

Methods of generating slide shows that use an image including a face are provided in accordance with the generation of one or more new images. A group of pixels is identified that correspond to a face within a digitally-acquired image. A zoom portion of the image including

the group of pixels may be determined. The image may be automatically zoomed to generate a zoomed image including the face enlarged by the zooming, or an option to generate the zoomed image may be provided. A center point of zooming in or out and an amount of zooming in or out may be determined after which another image may be automatically generated including a zoomed version of the face, or an option to generate the image including the zoomed version of the face may be provided. One or more new images may be generated each including a new group of pixels corresponding to the face, automatic panning may be provided using the one or more new images.

A method of generating one or more new digital images using an original digitally-acquired image including a face is further provided. One or more groups of pixels may be identified that correspond to two or more faces within the original digitally-acquired image. A portion of the original image may be selected to include the group of pixels. Values of pixesl may be automatically generated of one or more new images based on the selected portion in a manner which always includes at least one of the two or more faces within the one or more new images or a panning intermediate image between two of the faces of said two or more identified faces or a combination thereof.

Panning may be performed between the two or more identified faces. The panning may be from a first face to a second face of the two or more identified faces, and the second face may then be zoomed. The first face may be de-zoomed prior to panning to the second face. The second face may also be zoomed. The panning may include identifying a panning direction parameter between two of the identified faces. The panning may include sequencing along the identified panning direction between the two identified faces according to the identified panning direction parameter.

A method of generating a simulated camera movement in a still image using an original digitally acquired image including a face or other image feature is further provided. Simulated camera movements such as panning, tilting and zooming may be determined based on the orientation of the face or multiple faces or other features in an image to simulate the direction of the face and in particular the eyes. Such movement may then simulate the direction the

photographed subject is looking at. Such method may be extended to two or more identified faces, or as indicated other image features.

Each of the methods provided are preferably implemented within software and/or firmware either in the camera or with external processing equipment. The software may also be downloaded into the camera or image processing equipment. In this sense, one or more processor readable storage devices having processor readable code embodied thereon are provided. The processor readable code programs one or more processors to perform any of the above or below described methods.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1a illustrates a preferred embodiment of the main workflow of correcting images based on finding faces in the images.

Figure 1b illustrates a generic workflow of utilizing face information in an image to adjust image acquisition parameters in accordance with a preferred embodiment.

Figure 1c illustrates a generic workflow of utilizing face information in a single or a plurality of images to adjust the image rendering parameters prior to outputting the image in accordance with a preferred embodiment.

Figures 2a-2e illustrate image orientation based on orientation of faces in accordance with one or more preferred embodiments.

Figures 3a-3f illustrate an automatic composition and cropping of an image based on the location of the face in accordance with one or more preferred embodiments.

Figures 4a-4g illustrate digital fill-flash in accordance with one or more preferred embodiments.

Figure 4h describes an illustrative system in accordance with a preferred embodiment to determine in the camera as part of the acquisition process, whether fill flash is needed, and of so, activate such flash when acquiring the image based on the exposure on the face

Figure 5 illustrates the use of face-detection for generating dynamic slide shows, by applying automated and suggested zooming and panning functionality where the decision as to the center of the zoom is based on the detection of faces in the image.

Figure 6 describes an illustrative simulation of a viewfinder in a video camera or a digital camera with video capability, with an automatic zooming and tracking of a face as part of the live acquisition in a video camera, in accordance with a preferred embodiment.

Figure 7a and 7b illustrate an automatic focusing capability in the camera as part of the acquisition process based on the detection of a face in accordance with one or more preferred embodiments.

Figure 8 illustrates an adjustable compression rate based on the location of faces in the image in accordance with a preferred embodiment.

INCORPORATION BY REFERENCE

What follows is a cite list of references each of which is, in addition to that which is described as background, the invention summary, the abstract, the brief description of the drawings and the drawings themselves, hereby incorporated by reference into the detailed description of the preferred embodiments below, as disclosing alternative embodiments of elements or features of the preferred embodiments not otherwise set forth in detail below. A single one or a combination of two or more of these references may be consulted to obtain a variation of the preferred embodiments described in the detailed description herein:

United States patents no. RE33682, RE31370, 4,047,187, 4,317,991, 4,367,027, 4,638,364, 5,291,234, 5,432,863, 5,488,429, 5,638,136, 5,710,833, 5,724,456, 5,751,836, 5,781,650, 5,812,193, 5,818,975, 5,835,616, 5,870,138, 5,978,519, 5,991,456, 6,097,470, 6,101,271, 6,128,397, 6,134,339, 6,148,092, 6,151,073, 6,188,777, 6,192,149, 6,249,315, 6,263,113, 6,268,939, 6,278,491, 6,282,317, 6,301,370, 6,332,033, 6,393,148, 6,404,900, 6,407,777, 6,421,468, 6,438,264, 6,456,732, 6,459,436, 6,473,199, 6,501,857, 6,504,942, 6,504,951, 6,516,154, and 6,526,161;

United States published patent applications no. 2005/0041121, 2004/0114796, 2004/0240747, 2004/0184670, 2003/0071908, 2003/0052991, 2003/0044070, 2003/0025812, 2002/0172419, 2002/0136450, 2002/0114535, 2002/0105662, and 2001/0031142;

Published PCT applications no. WO 03/071484 and WO 02/045003

European patent application no EP 1 429 290 A;

Japanese patent application no. JP5260360A2;

British patent application no. GB0031423.7;

Yang et al., IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, no. 1, pp 34-58 (Jan. 2002);

Baluja & Rowley , "Neural Network-Based Face Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 1, pages 23-28, January 1998; and

Joffe, S. Ed, Institute of Electrical and Electronics Engineering, Red Eye Detection with Machine Learning, Proceedings 2003 International Conference of Image Processing. ICIP-2003. Barcelona, Spain, Sept. 14-17, 2003, New York, NY: IEEE, US, vol. 2 or 3, 14 September 2003, pages 871-874.

ILLUSTRATIVE DEFINITIONS

"Face Detection" involves the art of isolating and detecting faces in a digital image; Face Detection includes a process of determining whether a human face is present in an input image, and may include or is preferably used in combination with determining a position and/or other features, properties, parameters or values of parameters of the face within the input image;

"Image-enhancement" or "image correction" involves the art of modifying a digital image to improve its quality; such modifications may be "global" applied to the entire image , or "selective" when applied differently to different portions of the image. Some main categories non-exhaustively include: (i) Contrast Normalization and Image Sharpening.

(ii) Image Crop, Zoom and Rotate.

(iii) Image Color Adjustment and Tone Scaling.

(iv) Exposure Adjustment and Digital Fill Flash applied to a Digital Image.

(v) Brightness Adjustment with Color Space Matching; and Auto-Gamma determination with Image Enhancement.

(vi) Input/Output device characterizations to determine Automatic/Batch Image Enhancements.

(vii) In-Camera Image Enhancement

(viii) Face Based Image Enhancement

"Auto-focusing" involves the ability to automatically detect and bring a photographed object into the focus field;

"Fill Flash" involves a method of combining available light, such as sun light with another light source such as a camera flash unit in such a manner that the objects close to the camera, which may be in the shadow, will get additional exposure using the flash unit.

A "pixel" is a picture element or a basic unit of the composition of a digital image or any of the small discrete elements that together constitute an image;

"Digitally-Captured Image" includes an image that is digitally located and held in a detector;

"Digitally-Acquired Image" includes an image that is digitally recorded in a permanent file and/or preserved in a more or less permanent digital form; and

"Digitally-Detected Image": an image comprising digitally detected electromagnetic waves.

"Digital Rendering Device": A digital device that renders digital encoded information such as pixels onto a different device. Most common rendering techniques include the conversion of digital data into hard copy such as printers, and in particular laser printers, ink jet printers or thermal printers, or soft copy devices such as monitors, television, liquid crystal display, LEDs, OLED, etc.

'Simulated camera movement" is defined as follows: given an image of a certain dimension (e.g. MxN) , a window which is a partial image is created out of the original image (of smaller dimension to the original image). By moving this window around the image, a simulated

camera movement is generated. The movement can be horizontal, also referred to as "panning", vertical also referred to as "tilt", or orthogonal to the image plane also referred to as "zooming, or a combination thereof. The simulated camera movement may also include the gradual selection of non-rectangular window, e.g., in the shape of a trapezoid, or changing rectangular dimensions, which can simulate changes in the perspective to simulate physical movement of the camera also referred to as "dolly". Thus, simulated camera movement can include any geometrical distortion and may create a foreshortening effect based on the location of the foreground and the background relative to the camera.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Preferred embodiments are described below including methods and devices for providing or suggesting options for automatic digital image enhancements based on information relating to the location, position, focus, exposure or other parameter or values of parameters of region of interests and in particular faces in an image. Such parameters or values of parameter may include a spatial parameter.

A still image may be animated and used in a slide show by simulated camera movement, e.g., zooming, panning and/or rotating where the center point of an image is within a face or at least the face is included in all or substantially all of the images in the slide show.

A preferred embodiment includes an image processing application whether implemented in software or in firmware, as part of the image capture process, image rendering process, or as part of post processing. This system receives images in digital form, where the images can be translated into a grid representation including multiple pixels. This application detects and isolates the faces from the rest of the picture, and determines sizes and locations of the faces relative to other portions of the image or the entire image. Orientations of the faces may also be determined. Based on information regarding detected faces, preferably separate modules of the system collect facial data and perform image enhancement operations based on the collected

facial data. Such enhancements or corrections include automatic orientation of the image, color correction and enhancement, digital fill flash simulation and dynamic compression.

Advantages of the preferred embodiments include the ability to automatically perform or suggest or assist in performing complex tasks that may otherwise call for manual intervention and/or experimenting. Another advantage is that important regions, e.g., faces, of an image may be assigned, marked and/or mapped and then processing may be automatically performed and/or suggested based on this information relating to important regions of the images. Automatic assistance may be provided to a photographer in the post processing stage. Assistance may be provided to the photographer in determining a focus and an exposure while taking a picture. Meta-data may be generated in the camera that would allow an image to be enhanced based on the face information.

Many advantageous techniques are provided in accordance with preferred and alternative embodiments set forth herein. For example, a method of processing a digital image using face detection within said image to achieve one or more desired image processing parameters is provided. A group of pixels is identified that correspond to an image of a face within the digital image. Default values are determined of one or more parameters of at least some portion of said digital image. Values of the one or more parameters are adjusted within the digitally-detected image based upon an analysis of said digital image including said image of said face and said default values.

The digital image may be digitally-acquired and/or may be digitally-captured. Decisions for processing the digital image based on said face detection, selecting one or more parameters and/or for adjusting values of one or more parameters within the digital image may be automatically, semi-automatically or manually performed. Similarly, on the other end of the image processing workflow, the digital image may be rendered from its binary display onto a print, or a electronic display.

One or more different degrees of simulated fill flash may be created by manual, semi-automatic or automatic adjustment. The analysis of the image of the face may include a comparison of an overall exposure to an exposure around the identified face. The exposure may

be calculated based on a histogram. Digitally simulation of a fill flash may include optionally adjusting tone reproduction and/or locally adjusting sharpness. One or more objects estimated to be closer to the camera or of higher importance may be operated on in the simulated fill-flash. These objects determined to be closer to the camera or of higher importance may include one or more identified faces. A fill flash or an option for providing a suggested fill-flash may be automatically provided. The method may be performed within a digital acquisition device, a digital rendering device, or an external device or a combination thereof.

The face pixels may be identified, a false indication of another face within the image may be removed, and an indication of a face may be added within the image, each manually by a user, or semi-automatically or automatically using image processing apparatus. The face pixels identifying may be automatically performed by an image processing apparatus, and a manual verification of a correct detection of at least one face within the image may be provided.

A method of digital image processing using face detection to achieve a desired image parameter is further provided including identifying a group of pixels that correspond to an image of a face within a digitally-detected image. Initial values of one or more parameters of at least some of the pixels are determined. An initial parameter of the digitally-detected image is determined based on the initial values. Values of the one or more parameters of pixels within the digitally-detected image are automatically adjusted based upon a comparison of the initial parameter with the desired parameter or an option for adjusting the values is automatically provided.

The digitally-detected image may include a digitally-acquired, rendered and/or digitally-captured image. The initial parameter of the digitally-detected image may include an initial parameter of the face image. The one or more parameters may include any of orientation, color, tone, size, luminance, and focus. The method may be performed within a digital camera as part of a pre-acquisition stage, within a camera as part of post processing of the captured image or within external processing equipment. The method may be performed within a digital rendering device such as a printer, or as a preparation for sending an image to an output device, such as in the print driver, which may be located in the printer or on an external device such as the PC, as

part of a preparation stage prior to displaying or printing the image. An option to manually remove a false indication of a face or to add an indication of a face within the image may be included. An option to manually override, the automated suggestion of the system, whether or not faces were detected, may be included.

The method may include identifying one or more sub-groups of pixels that correspond to one or more facial features of the face. Initial values of one or more parameters of pixels of the one or more sub-groups of pixels may be determined. An initial spatial parameter of the face within the digital image may be determined based on the initial values. The initial spatial parameter may include any of orientation, size and location.

When the spatial parameter is orientation, values of one or more parameters of pixels may be adjusted for re-orienting the image to an adjusted orientation. The one or more facial features may include one or more of an eye, a mouth, two eyes, a nose, an ear, neck, shoulders and/or other facial or personal features, or other features associated with a person such as an article of clothing, furniture, transportation, outdoor environment (e.g., horizon, trees, water, etc.) or indoor environment (doorways, hallways, ceilings, floors, walls, etc.), wherein such features may be indicative of an orientation. The one or more facial or other features may include two or more features, and the initial orientation may be determined base on relative positions of the features that are determined based on the initial values. A shape such as a triangle may be generated for example between the two eyes and the center of the mouth, a golden rectangle as described above, or more generically, a polygon having points corresponding to preferably three or more features as vertices or axis.

Initial values of one or more chromatic parameters, such as color and tone, of pixels of the digital image may be determined. The values of one or more parameters may be automatically adjusted or an option to adjust the values to suggested values may be provided.

The method may be performed within any digital image capture device, which as, but not limited to digital still camera or digital video camera. The one or more parameters may include overall exposure, relative exposure, orientation, color balance, white point, tone reproduction, size, or focus, or combinations thereof. The face pixels identifying may be automatically

performed by an image processing apparatus, and the method may include manually removing one or more of the groups of pixels that correspond to an image of a face. An automatically detected face may be removed in response to false detection of regions as faces, or in response to a determination to concentrate on less image faces or images faces that were manually determined to be of higher subjective significance, than faces identified in the identifying step. A face may be removed by increasing a sensitivity level of said face identifying step. The face removal may be performed by an interactive visual method, and may use an image acquisition built-in display.

The face pixels identifying may be performed with an image processing apparatus, and may include manually adding an indication of another face within the image. The image processing apparatus may receive a relative value as to a detection assurance or an estimated importance of the detected regions. The relative value may be manually modified as to the estimated importance of the detected regions.

Within a digital camera, a method of digital image processing using face detection for achieving a desired image parameter is further provided. A group of pixels is identified that correspond to a face within a digital image. First initial values of a parameter of pixels of the group of pixels are determined, and second initial values of a parameter of pixels other than pixels of the group of pixels are also determined. The first and second initial values are compared. Adjusted values of the parameter are determined based on the comparing of the first and second initial values and on a comparison of the parameter corresponding to at least one of the first and second initial values and the desired image parameter.

Initial values of luminance of pixels of the group of pixels corresponding to the face may be determined. Other initial values of luminance of pixels other than the pixels corresponding to the face may also be determined. The values may then be compared, and properties of aperture, shutter, sensitivity and a fill flash may be determined for providing adjusted values corresponding to at least some of the initial values for generating an adjusted digital image. The pixels corresponding to the face may be determined according to sub-groups corresponding to one or more facial features.

A method of generating one or more new digital images using an original digitally-acquired image including a face is further provided. A group of pixels that correspond to a face within the original digitally-acquired image is identified. A portion of the original image is selected to include the group of pixels. Values of pixels of one or more new images based on the selected portion are automatically generated, or an option to generate them is provided, in a manner which always includes the face within the one or more new images.

A transformation may be gradually displayed between the original digitally-acquired image and one or more new images. Parameters of said transformation may be adjusted between the original digitally-acquired image and one or more new images. Parameters of the transformation between the original digitally-acquired image, e.g., including aface, and one or more new images may be selected from a set of at least one or more criteria including timing or blending or a combination thereof. The blending may include dissolving, flying, swirling, appearing, flashing, or screening, or combinations thereof.

Methods of generating slide shows that use an image including a face are provided in accordance with the generation of one or more new images. A group of pixels is identified that correspond to a face within a digitally-acquired image. A zoom portion of the image including the group of pixels may be determined. The image may be automatically zoomed to generate a zoomed image including the face enlarged by the zooming, or an option to generate the zoomed image may be provided. A center point of zooming in or out and an amount of zooming in or out may be determined after which another image may be automatically generated including a zoomed version of the face, or an option to generate the image including the zoomed version of the face may be provided. One or more new images may be generated each including a new group of pixels corresponding to the face, automatic panning may be provided using the one or more new images.

A method of generating one or more new digital images using an original digitally-acquired image including a face is further provided. One or more groups of pixels may be identified that correspond to two or more faces within the original digitally-acquired image. A portion of the original image may be selected to include the group of pixels. Values of pixels

may be automatically generated of one or more new images based on the selected portion in a manner which always includes at least one of the two or more faces within the one or more new images or a panning intermediate image between two of the faces of said two or more identified faces or a combination thereof.

Panning may be performed between the two or more identified faces. The panning may be from a first face to a second face of the two or more identified faces, and the second face may then be zoomed. The first face may be de-zoomed prior to panning to the second face. The second face may also be zoomed. The panning may include identifying a panning direction parameter between two of the identified faces. The panning may include sequencing along the identified panning direction between the two identified faces according to the identified panning direction parameter.

A method of digital image processing using face detection for achieving a desired spatial parameter is further provided including identifying a group of pixels that correspond to a face within a digital image, identifying one or more sub-groups of pixels that correspond to one or more facial features of the face, determining initial values of one or more parameters of pixels of the one or more sub-groups of pixels, determining an initial spatial parameter of the face within the digital image based on the initial values, and determining adjusted values of pixels within the digital image for adjusting the image based on a comparison of the initial and desired spatial parameters.

The initial spatial parameter may include orientation. The values of the pixels may be automatically adjusted within the digital image to adjust the initial spatial parameter approximately to the desired spatial parameter. An option may be automatically provided for adjusting the values of the pixels within the digital image to adjust the initial spatial parameter to the desired spatial parameter.

A method of digital image processing using face detection is also provided wherein a first group of pixels that correspond to a face within a digital image is identified, and a second group of pixels that correspond to another feature within the digital image is identified. A re-compositioned image is determined including a new group of pixels for at least one of the face

and the other feature. The other feature may include a second face. The re-compositioned image may be automatically generated or an option to generate the re-compositioned image may be provided. Values of one or more parameters of the first and second groups of pixels, and relative-adjusted values, may be determined for generating the re-compositioned image.

Each of the methods provided are preferably implemented within software and/or firmware either in the camera, the rendering device such as printers or display, or with external processing equipment. The software may also be downloaded into the camera or image processing equipment. In this sense, one or more processor readable storage devices having processor readable code embodied thereon are provided. The processor readable code programs one or more processors to perform any of the above or below described methods.

Figure 1a illustrates a preferred embodiment. An image is opened by the application in block 102. The software then determines whether faces are in the picture as described in block 106. If no faces are detected, the software ceases to operate on the image and exits, 110.

Alternatively, the software may also offer a manual mode, where the user, in block 116 may inform the software of the existence of faces, and manually marks them in block 118. The manual selection may be activated automatically if no faces are found, 116, or it may even be optionally activated after the automatic stage to let the user, via some user interface to either add more faces to the automatic selection 112 or even 114, remove regions that are mistakenly 110 identified by the automatic process 118 as faces. Additionally, the user may manually select an option that invokes the process as defined in 106. This option is useful for cases where the user may manually decide that the image can be enhanced or corrected based on the detection of the faces. Various ways that the faces may be marked, whether automatically of manually, whether in the camera or by the applications, and whether the command to seek the faces in the image is done manually or automatically, are all included in preferred embodiments herein.

In an alternative embodiment, the face detection software may be activated inside the camera as part of the acquisition process, as described in Block 104. This embodiment is further depicted in Figure 1b. In this scenario, the face detection portion 106 may be implemented differently to support real time or near real time operation. Such implementation may include

sub-sampling of the image, and weighted sampling to reduce the number of pixels on which the computations are performed.

In an alternative embodiment, the face detection software may be activated inside the rendering device as part of the output process, as described in Block 103. This embodiment is further depicted in Figure 1c. In this scenario, the face detection portion 106 may be implemented either within the rendering device, or within a en external driver to such device.

After the faces are tagged, or marked, whether manually as defined in 106, or automatically, 118, the software is ready to operate on the image based on the information generated by the face-detection stage. The tools can be implemented as part of the acquisition, as part of the post-processing, or both.

Block 120 describes panning and zooming into the faces. This tool can be part of the acquisition process to help track the faces and create a pleasant composition, or as a post processing stage for either cropping an image or creating a slide show with the image, which includes movement. This tool is further described in Figure 6.

Block 130 depicts the automatic orientation of the image, a tool that can be implemented either in the camera as art of the acquisition post processing, or on a host software. This tool is further described in Figures 2a-2e.

Block 140 describes the way to color-correct the image based on the skin tones of the faces. This tool can be part of the automatic color transformations that occur in the camera when converting the image from the RAW sensor data form onto a known, e.g. RGB representation, or later in the host, as part of image enhancement software. The various image enhancement operations may be global, affecting the entire image, such as rotation, and/or may be selective based on local criteria. For example, in a selective color or exposure correction as defined in block 140, a preferred embodiment includes corrections done to the entire image, or only to the face regions in a spatially masked operation, or to specific exposure, which is a luminance masked operation. Note also that such masks may include varying strength, which correlates to varying degrees of applying a correction. This allows a local enhancement to better blend into the image.

Block 150 describes the proposed composition such as cropping and zooming of an image to create a more pleasing composition. This tool, 150 is different from the one described in block 120 where the faces are anchors for either tracking the subject or providing camera movement based on the face location.

Block 160 describes the digital-fill-flash simulation which can be done in the camera or as a post processing stage. This tool is further described in Figures 4a-4e.
Alternatively to the digital fill flash, this tool may also be an actual flash sensor to determine if a fill flash is needed in the overall exposure as described in Block 170. In this case, after determining the overall exposure of the image, if the detected faces in the image are in the shadow, a fill flash will automatically be used. Note that the exact power of the fill flash, which should not necessarily be the maximum power of the flash, may be calculated based on the exposure difference between the overall image and the faces. Such calculation is well known to the one skilled in the art and is based on a tradeoff between aperture, exposure time, gain and flash power.

This tool is further described in Figure 4e. Block 180 describes the ability of the camera to focus on the faces. This can be used as a pre-acquisition focusing tool in the camera, as further illustrated in Figure 7.

Referring to Figure 1b, which describes a process of using face detection to improve in camera acquisition parameters, as aforementioned in Figure 1a, block 106. In this scenario, a camera is activated, 1000, for example by means of half pressing the shutter, turning on the camera, etc. The camera then goes through the normal pre-acquisition stage to determine, 1004, the correct acquisition parameters such as aperture, shutter speed, flash power, gain, color balance, white point, or focus. In addition, a default set of image attributes, particularly related to potential faces in the image, are loaded, 1002. Such attributes can be the overall color balance, exposure, contrast, orientation etc.

An image is then digitally captured onto the sensor, 1010. Such action may be continuously updated, and may or may not include saving such captured image into permanent storage.

An image-detection process, preferably a face detection process, is applied to the captured image to seek faces in the image, 1020. If no images are found, the process terminates, 1032. Alternatively, or in addition to the automatic detection of 1030, the user can manually select, 1034 detected faces, using some interactive user interface mechanism, by utilizing, for example, a camera display. Alternatively, the process can be implemented without a visual user interface by changing the sensitivity or threshold of the detection process.

When faces are detected, 1040, they are marked, and labeled. Detecting defined in 1040 may be more than a binary process of selecting whether a face is detected or not, It may also be designed as part of a process where each of the faces is given a weight based on size of the faces, location within the frame, other parameters described herein, etc., which define the importance of the face in relation to other faces detected.

Alternatively, or in addition, the user can manually deselect regions, 1044 that were wrongly false detected as faces. Such selection can be due to the fact that a face was false detected or when the photographer may wish to concentrate on one of the faces as the main subject matter and not on other faces. Alternatively, 1046, the user may re-select, or empahsize one or more faces to indicate that these faces have a higher importance in the calculation relative to other faces. This process as defined in 1046, further defines the preferred identification process to be a continuous value one as opposed to a binary one. The process can be done utilizing a visual user interface or by adjusting the sensitivity of the detection process.
After the faces are correctly isolated, 1040, their attributes are compared, 1050 to default values that were predefined in 1002. Such comparison will determine a potential transformation between the two images, in order to reach the same values. The transformation is then translated to the camera capture parameters, 1070, and the image, 1090 is acquired.

A practical example is that if the captured face is too dark, the acquisition parameters may change to allow a longer exposure, or open the aperture. Note that the image attributes are not necessarily only related to the face regions but can also be in relations to the overall exposure. As an exemplification, if the overall exposure is correct but the faces are

underexposed, the camera may shift into a fill-flash mode as subsequently illustrated in Figure 4a-4f.

Figure 1c illustrates a process of using face detection to improve output or rendering parameters, as aforementioned in Figure 1a, block 103. In this scenario, a rendering device such as a printer or a display, herein referred to as the Device, activated, 1100. Such activation can be performed for example within a printer, or alternatively within a device connected to the printer such as a PC or a camera. The device then goes through the normal pre-rendering stage to determine, 1104, the correct rendering parameters such as tone reproduction, color transformation profiles, gain, color balance, white point and resolution. In addition, a default set of image attributes, particularly related to potential faces in the image, are loaded, 1102. Such attributes can be the overall color balance, exposure, contrast, orientation etc.

An image is then digitally downloaded onto the device, 1110. An image-detection process, preferably a face detection process, is applied to the downloaded image to seek faces in the image, 1120. If no images are found, the process terminates, 1132 and the device resumes its normal rendering process. Alternatively, or in addition to the automatic detection of 1130, the user can manually select, 1134 detected faces, using some interactive user interface mechanism, by utilizing, for example, a display on the device. Alternatively, the process can be implemented without a visual user interface by changing the sensitivity or threshold of the detection process. When faces are detected, 1040, they are marked, and labeled. Detecting defined in 1140 may be more than a binary process of selecting whether a face is detected or not, It may also be designed as part of a process where each of the faces is given a weight based on size of the faces, location within the frame, other parameters described herein, etc., which define the importance of the face in relation to other faces detected.

Alternatively, or in addition, the user can manually deselect regions, 1144 that were wrongly false detected as faces. Such selection can be due to the fact that a face was false detected or when the photographer may wish to concentrate on one of the faces as the main subject matter and not on other faces. Alternatively, 1146, the user may re-select, or emphasize one or more faces to indicate that these faces have a higher importance in the calculation relative

to other faces. This process as defined in 1146, further defines the preferred identification process to be a continuous value one as opposed to a binary one. The process can be done utilizing a visual user interface or by adjusting the sensitivity of the detection process. After the faces are correctly isolated, 1140, their attributes are compared, 1150 to default values that were predefined in 1102. Such comparison will determine a potential transformation between the two images, in order to reach the same values. The transformation is then translated to the device rendering parameters, 1170, and the image, 1190 is rendered. The process may include a plurality of images. In this case 1180, the process repeats itself for each image prior to performing the rendering process. A practical example is the creation of a thumbnail or contact sheet whish is a collection of low resolution images, on a single display instance.

A practical example is that if the face was too dark captured, the rendering parameters may change the tone reproduction curve to lighten the face. Note that the image attributes are not necessarily only related to the face regions but can also be in relations to the overall tone reproduction.

Referring to Figures 2a-2e, which describe the invention of automatic rotation of the image based on the location and orientation of faces, as highlighted in Figure-1 Block 130. An image of two faces is provided in Figure 2a. Note that the faces may not be identically oriented, and that the faces may be occluding.

The software in the face detection stage, including the functionality of Figure 1a, blocks 108 and 118, will mark the two faces, of the mother and son as an estimation of an ellipse 210 and 220 respectively. Using known mathematical means, such as the covariance matrix of the ellipse, the software will determine the main axis of the two faces, 212 and 222 respectively as well as the secondary axis 214 and 224. Even at this stage, by merely comparing the sizes of the axis, the software may assume that the image is oriented 90 degrees, in the case that the camera hel helo in landscape mode, which is horizontal, or in portrait mode which is vertical or +90 degrees, aka clockwise, or –90 degrees aka counter clockwise. Alternatively, the application may also be utilized for any arbitrary rotation value. However this information may not suffice to decide whether the image is rotated clockwise or counter-clockwise.

Figure 2c describes the step of extracting the pertinent features of a face, which are usually highly detectable. Such objects may include the eyes, 214, 216 and 224, 226, and the lips, 218 and 228. The combination of the two eyes and the center of the lips creates a triangle 230 which can be detected not only to determine the orientation of the face but also the rotation of the face relative to a facial shot. Note that there are other highly detectable portions of the image which can be labeled and used for orientation detection, such as the nostrils, the eyebrows, the hair line, nose bridge and the neck as the physical extension of the face etc. In this figure, the eyes and lips are provided as an example of such facial features Based on the location of the eyes if found, and the mouth, the image may, e.g., need to be rotated in a counter clockwise direction.

Note that it may not be enough to just locate the different facial features, but it may be necessary to compare such features to each other. For example, the color of the eyes may be compared to ensure that the pair of eyes originate form the same person. Another example is that in figures 2-c and 2-d, if the software combined the mouth of 218 with the eyes of 226, 224, the orientation would have been determined as clockwise. In this case, the software detects the correct orientation by comparing the relative size of the mouth and the eyes. The above method describes means of determining the orientation of the image based on the relative location of the different facial objects. For example, it may be desired that the two eyes should be horizontally situated, the nose line perpendicular to the eyes, the mouth under the nose etc. Alternatively, orientation may be determined based on the geometry of the facial components themselves. For example, it may be desired that the eyes are elongated horizontally, which means that when fitting an ellipse on the eye, such as described in blocs 214 and 216, it may be desired that the main axis should be horizontal. Similar with the lips which when fitted to an ellipse the main axis should be horizontal. Alternatively, the region around the face may also be considered. In particular, the neck and shoulders which are the only contiguous skin tone connected to the head can be an indication of the orientation and detection of the face.

Figure 2-e illustrates the image as correctly oriented based on the facial features as detected. In some cases not all faces will be oriented the same way. In such cases, the software

may decide on other criteria to determine the orientation of the prominent face in the image. Such determination of prominence can be based on the relevant size of the faces, the exposure, or occlusion.

If a few criteria are tested, such as the relationship between different facial components and or the orientation of individual components, not all results will be conclusive to a single orientation. This can be due to false detections, miscalculations, occluding portions of faces, including the neck and shoulders, or the variability between faces. In such cases, a statistical decision may be implemented to address the different results and to determine the most likely orientation. Such statistical process may be finding the largest results (simple count), or more sophisticated ordering statistics such as correlation or principal component analysis, where the basis function will be the orientation angle. Alternatively or in addition, the user may manually select the prominent face or the face to be oriented. The particular orientation of the selected or calculated prominent face may itself be automatically determined, programmed, or manually determined by a user.

The process for determining the orientation of images can be implemented in a preferred embodiment as part of a digital display device. Alternatively, this process can be implemented as part of a digital printing device, or within a digital acquisition device.

The process can also be implemented as part of a display of multiple images on the same page or screen such as in the display of a contact-sheet or a thumbnail view of images. In this case, the user may approve or reject the proposed orientation of the images individually or by selecting multiple images at once. In the case of a sequence of images, this invention may also determine the orientation of images based on the information as approved by the user regarding previous images.

Figures 3a-3f describe an illustrative process in which a proposed composition is offered based on the location of the face. As defined in Figure 1a blocks 108 and 118, the face 320 is detected as are one or more pertinent features, as illustrated in this case, the eyes 322 and 324. The location of the eyes are then calculated based on the horizontal, 330 and vertical 340

location. In this case, the face is located at the center of the image horizontally and at the top quarter vertically as illustrated in Figure 3-d.

Based on common rules of composition and aesthetics, e.g., a face in a close up may be considered to be better positioned, as in Figure 3-e if the eyes are at the 2/3rd line as depicted in 350, and 1/3 to the left or 1/3 to the right as illustrated in 360. Other similar rules may be the location of the entire face and the location of various portions of the face such as the eyes and lips based on aesthetic criteria such as the applying the golden-ratio for faces and various parts of the face within an image.

Figure 3c introduces another aspect of face detection which may happen especially in non-restrictive photography. The faces may not necessarily be frontally aligned with the focal plane of the camera. In this figure, the object is looking to the side exposing partial frontal, or partial profile of the face. In such cases, the software may elect to use, the center of the face, which in this case may align with the left eye of the subject. If the subject was in full frontal position, the software may determine the center of the face to be around the nose bridge. The center of the face may be determined to be at the center of a rectangle, ellipse or other shape generally determined to outline the face or at the intersection of cross-hairs or otherwise as may be understood by those skilled in the art (see, e.g., ellipse 210 of Figures 2b-2e, ellipse 320 of Figure 3b, ellipse 330 of Figure 3c, the cross-hairs 350, 360 of Figure 3e).

Based on the knowledge of the face and its pertinent features such as eyes, lips nose and ears, the software can either automatically or via a user interface that would recommend the next action to the user, crop portions of the image to reach such composition. For this specific image, the software will eliminate the bottom region 370 and the right portion 380. The process of re-compositioning a picture is subjective. In such case this invention will act as guidance or assistance to the user in determining the most pleasing option out of potentially a few. In such a case a plurality of proposed compositions can be displayed and offered to the user and the user will select one of them.

In an alternative embodiment, the process of re-compositioning the image can be performed within the image acquisition device as part of the image taking process, whether as a

pre-capture, pre-acquisition or post acquisition stage. In this scenario the acquisition device may display a proposed re-compositioning of the image on its display. Such re-compositioning may be displayed in the device viewfinder or display similarly to Figure 3f, or alternatively as guidelines of cropping such as lines 352 and 354. A user interface such will enable the user to select form the original composed image, or the suggested one. Similar functionality can be offered as part of the post acquisition or otherwise referred to the playback mode.

In additional embodiments, the actual lines of aesthetics, for example, the $1/3^{rd}$ lines 350 and 350, may also be displayed to the use as assistance in determining the right composition. Referring to Figures 4a-4f, the knowledge of the faces may assist the user in creating an automatic effect that is otherwise created by a fill-flash. Fill-flash is a flash used where the main illumination is available light. In this case, the flash assists in opening up shadows in the image. Particularly, fill flash is used for images where the object in the foreground is in the shadow. Such instances occur for example when the sun is in front of the camera, thus casting a shadow on the object in the foreground. In many cases the object includes people posing in front of a background of landscape.

Figure 4a illustrates such image. The overall image is bright due to the reflection of the sun in the water. The individuals in the foreground are therefore in the shadow.

A certain embodiment of calculating the overall exposure can be done using an exposure histogram. Those familiar in the art may decide on other means of determining exposure, any of which may be used in accordance with an alternative embodiment. When looking at the histogram of the luminance of the image at Figure 4-b, there are three distinct areas of exposure which correspond to various areas. The histogram depicts the concentration of pixels, as defined by the Y-Axis 416, as a function of the different gray levels as defined by the X-axis 418. The higher the pixel count for a specific gray level, the higher the number as depicted on the y-axis. Regions 410 are in the shadows which belong primarily to the mother. The midtones in area 412 belong primarily to the shaded foreground water and the baby. The highlights 414 are the water. However, not all shadows may be in the foreground, and not all highlights may be in the

background. A correction of the exposure based on the histogram may result in an unnatural correction.

When applying face detection, as depicted in Figure 4-c, the histogram in Figure 4-d may be substantially more clear. In this histogram, region 440 depicts the faces which are in the shadow. Note that the actual selection of the faces, as illustrated in 4-c need not be a binary mask but can be a gray scale mask where the boundaries are feathered or gradually changing. In addition, although somewhat similar in shape, the face region 440 may not be identical to the shadow region of the entire image, as defined, e.g., in Figure 4b at area 410. By applying exposure correction to the face regions as illustrated in figure 4-e, such as passing the image through a lookup table 4-f, the effect is similar to the one of a fill flash that illuminated the foreground, but did not affect the background. By taking advantage of the gradual feathered mask around the face, such correction will not be accentuated and noticed.
Figure 4e can also be performed manually thus allowing the user to create a varying effect of simulated fill flash. Alternatively, the software may present the user with a selection of corrections based on different tone reproduction curves and different regions for the user to choose from.

Although exposure, or tone reproduction, may be the most preferred enhancement to simulate fill flash, other corrections may apply such as sharpening of the selected region, contrast enhancement, of even color correction. Additional advantageous corrections may be understood by those familiar with the effect of physical strobes on photographed images.

Alternatively, as described by the flow chart of Figure 4g, a similar method may be utilized in the pre-acquisition stage, to determine if a fill flash is needed or not. The concept of using a fill flash is based on the assumption that there are two types of light sources that illuminate the image: an available external or ambient light source, which is controlled by the gain, shutter speed and aperture, and a flash which is only controlled by the flash power and affected by the aperture. By modifying the aperture vs. the shutter speed, the camera can either enhance the effect of the flash or decrease it, while maintaining the overall exposure.

Referring now to Figure 4g, a digital image is provided at 450. A determination is made at 460 whether faces were found in the image. As will be seen below, this process can be applied to other image features or regions within a digital image, e.g., a region including a face and also its surroundings, or a portion of a face less than the entire face, such as the eyes or the mouth or the nose, or two of these, or a background or foreground region within an image. If no faces (or other regions or features, hereinafter only "faces" will be referred to, as an example) are found, the process exits at 462. If a one or more faces is found at 460, then the faces are automatically marked at 464. There can be a manual step here instead of or in addition to the automatic marking at 464. A determination of exposure in face regions occurs at 470. Then, at 474 it is determined whether exposure of the face regions is lower than an overall exposure. If the exposure of the face regions is not lower than an overall exposure, then the image may be left as is by moving the process to 478. If the exposure of the face regions is lower than an overall exposure, then a fill flash may be digitally simulated at 480.

Referring still to Figure 4g, an exemplary digital fill flash simulation 480 includes creating masks to define one or more selected regions at 482a. Exposure of the selected regions is increased at 484a. Sharpening is applied to the selected regions at 486a. Tone reproduction is applied on selected regions 488a. Single or multiple results may be displayed to the user at 490a, and then a user selects a preferred results at 492a. An image may be displayed with a parameter to modify at 494a, and then a user adjusts the extent of modification at 496a. After 492a and/or 496a correction is applied to the image at 498.

Referring now to Figure 4h, when the user activates the camera, in block 104 (see also Figure 1a), the camera calculates the overall exposure, 482b. Such calculation is known to one skilled in the art and can be as sophisticated as needed. In block 108, the camera searched for the existence of faces in the image. An exposure is then calculated to the regions defined as belonging to the faces, 486b. The disparity between the overall exposure as determined in 484b and the faces, 486b is calculated. If the face regions are substantially darker than the overall exposure 486b, the camera will then activate the flash in a fill mode, 490b, calculate the necessary flash power, aperture and shutter speed, 492b and acquire the image 494b with the fill

flash. The relationship between the flash power, the aperture and the shutter speed are well formulated and known to one familiar in the art of photography. Examples of such calculations can be found in US patent 6,151,073 to Steinberg et. al., which is hereby incorporated by reference.

Alternatively, in a different embodiment, 496b, this algorithm may be used to simply determine the overall exposure based on the knowledge and the exposure of the faces. The image will then be taken, 488b, based on the best exposure for the faces, as calculated in 496b. Many cameras have a matrix type of exposure calculation where different regions receive different weights as to the contribution for the total exposure. In such cases, the camera can continue to implement the same exposure algorithm with the exception that now, regions with faces in them will receive a larger weight in their importance towards such calculations.

Figure 5 describes yet another valuable use of the knowledge of faces in images. In this example, knowledge of the faces can help improve the quality of image presentation. An image, 510 is inserted into slide show software. The face is then detected as defined in Figure 1 block 104, including the location of the important features of the face such as the eyes and the mouth.

The user can then choose between a few options such as: zoom into the face vs. zoom out of the face and the level of zoom for a tight close up 520, a regular close up 520 or a medium close up as illustrated by the bounding box 540. The software will then automatically calculate the necessary pan, tilt and zoom needed to smoothly and gradually switch between the beginning and the end state. In the case where more than one face is found, the software can also create a pan and zoom combination that will begin at one face and end at the other. In a more generic manner, the application can offer from within a selection of effects such as dissolve,

Figure 6 illustrates similar functionality but inside the device. A camera, whether still or video as illustrated by the viewfinder 610, when in auto track mode 600, can detect the faces in the image, and then propose a digital combination of zoom pan and tilt to move from the full wide image 630 to a zoomed in image 640. Such indication may also show on the viewfinder prior to zooming, 632 as indication to the user, which the user can then decide in real time whether to activate the auto zooming or not. This functionality can also be added to a tracking

mode where the camera continuously tracks the location of the face in the image. In addition, the camera can also maintain the right exposure and focus based on the face detection.

Figure 7a illustrates the ability to auto focus the camera based on the location of the faces in the image. Block 710 is a simulation of the image as seen in the camera viewfinder. When implementing a center weight style auto focus, 718, one can see that the image will focus on the grass, 17 feet away, as depicted by the cross 712. However, as described in this invention, if the camera in the pre-acquisition mode, 104 detects the face, 714, and focuses on the face, rather than arbitrarily on the center, the camera will then indicate to the user where the focus is, 722 and the lens will be adjusted to the distance to the face, which in this example, as seen in 728, is 11 ft. vs. the original 17 ft.

This process can be extended by one skilled in the art to support not only a single face, but multiple faces, by applying some weighted average. Such average will depend on the disparity between the faces, in distances, and sizes.

Figure 7b presents the workflow of the process as illustrated via the viewfinder in figure 7-a. When the face-auto-focus mode is activated, 740, the camera continuously seeks for faces, 750. This operation inside the camera is performed in real time and needs to be optimized as such. If no faces are detected 760, the camera will switch to an alternative focusing mode, 762. If faces are detected, the camera will mark the single or multiple faces. Alternatively, the camera may display the location of the face 772, on the viewfinder or LCD. The user may then take a picture, 790 where the faces are in focus.

Alternatively, the camera may shift automatically, via user request or through preference settings to a face-tracking mode 780. In this mode, the camera keeps track of the location of the face, and continuously adjusts the focus based on the location of the face.

In an alternative embodiment, the camera can search for the faces and mark them, similarly to the cross in figure 722. The photographer can then lock the focus on the subject, for example by half pressing the shutter. Locking the focus on the subject differs form locking the focus, by the fact that if the subject then moves, the camera can still maintain the correct focus by modifying the focus on the selected object.

Figure 8 describes the use of information about the location and size of faces to determine the relevant compression ratio of different segments of the image. An image 800 is segmented into tiles using horizontal grid 830 and vertical grid 820. The tiles which include or partially include face information are marked 850. Upon compression, regions of 850 may be compressed differently than the tiles of image 800 outside of this region. The degree of compression may be predetermined, pre-adjusted by the user or determined as an interactive process. In the case of multiple detected faces in an image, the user may also assign different quality values, or compression rates based on the importance of the faces in the image. Such importance may be determined subjectively using an interactive process, or objectively using parameters such as the relative size of the face, exposure or location of the face relative to other subjects in the image.

An alternative method of variable compression involves variable resolution of the image. Based on this, the method described with reference to Figure 8 can also be utilized to create variable resolution, where facial regions which are preferably usually the important regions of the image, and will be preferably maintained with higher overall resolution than other regions in the image. According to this method, referring to Figure 8, the regions of the face as defined in block 850 will be preferably maintained with higher resolution than regions in the image 800 which are not part of 850.

An image can be locally compressed so that specific regions will have a higher quality compression which equates to lower compression rate. Alternatively and/or correspondingly, specific regions of an image may have more or less information associated with them. The information can be encoded in a frequency-based, or temporal-based method such as JPEG or Wavelet encoding. Alternatively, compression on the spatial domain may also involve a change in the image resolution. Thus, local compression may also be achieved by defining adjustable variable resolution of an image in specific areas. By doing so, selected or determined regions of importance may maintain low compression or high resolution compared with regions determined to have less importance or non-selected regions in the image.

Face detection and face tracking technology, particularly for digital image processing applications according to preferred and alternative embodiments set forth herein, are further advantageous in accordance with various modifications of the systems and methods of the above description as may be understood by those skilled in the art, as set forth in the references cited and incorporated by reference herein and as may be otherwise described below. For example, such technology may be used for identification of faces in video sequences, particularly when the detection is to be performed in real-time. Electronic component circuitry and/or software or firmware may be included in accordance with one embodiment for detecting flesh-tone regions in a video signal, identifying human faces within the regions and utilizing this information to control exposure, gain settings, auto-focus and/or other parameters for a video camera (see, e.g., US patents 5,488,429 and 5,638,136 to Kojima et al., each hereby incorporated by reference). In another embodiment, a luminance signal and/or a color difference signal may be used to detect the flesh tone region in a video image and/or to generate a detecting signal to indicate the presence of a flesh tone region in the image. In a further embodiment, electronics and/or software or firmware may detect a face in a video signal and substitute a "stored" facial image at the same location in the video signal, which may be useful, e.g., in the implementation of a low-bandwidth videophone (see, e.g., US patent 5,870,138 to Smith et al., hereby incorporated by reference).

In accordance with another embodiment, a human face may be located within an image which is suited to real-time tracking of a human face in a video sequence (see, e.g., US patents 6,148,092 and 6,332,033 to Qian, hereby incorporated by reference). An image may be provided including a plurality of pixels and wherein a transformation and filtering of each pixel is performed to determine if a pixel has a color associated with human skin-tone. A statistical distribution of skin tones in two distinct directions may be computed and the location of a face within the image may be calculated from these two distributions.

In another embodiment, electrical and/or software or firmware components may be provided to track a human face in an image from a video sequence where there are multiple persons (see, e.g., US patent 6,404,900 also to Qian, hereby incorporated by reference). A

projection histogram of the filtered image may be used for output of the location and/or size of tracked faces within the filtered image. A face-like region in an image may also be detected by applying information to an observer tracking display of the auto-stereoscopic type (see, e.g., US patent 6,504,942 to Hong et al., incorporated by reference).

An apparatus according to another embodiment may be provided for detection and recognition of specific features in an image using an eigenvector approach to face detection (see, e.g., US patent 5,710,833 to Moghaddam et al., incorporated by reference). Additional eigenvectors may be used in addition to or alternatively to the principal eigenvector components, e.g., all eigenvectors may be used. The use of all eigenvectors may be intended to increase the accuracy of the apparatus to detect complex multi-featured objects.

Another approach may be based on object identification and recognition within a video image using model graphs and/or bunch graphs that may be particularly advantageous in recognizing a human face over a wide variety of pose angles (see, e.g., US patent 6,301,370 to Steffens et al., incorporated by reference). A further approach may be based on object identification, e.g., also using eigenvector techniques (see, e.g., US patent 6,501,857 to Gotsman et al., incorporated by reference). This approach may use smooth weak vectors to produce near-zero matches, or alternatively, a system may employ strong vector thresholds to detect matches. This technique may be advantageously applied to face detection and recognition in complex backgrounds.

Another field of application for face detection and/or tracking techniques, particularly for digital image processing in accordance with preferred and alternative embodiments herein, is the extraction of facial features to allow the collection of biometric data and tracking of personnel, or the classification of customers based on age, sex and other categories which can be related to data determined from facial features. Knowledge-based electronics and/or software or firmware may be used to provide automatic feature detection and age classification of human faces in digital images (see, e.g., US patent 5,781,650 to Lobo & Kwon, hereby incorporated by reference). Face detection and feature extraction may be based on templates (see US patent 5,835,616 also to Lobo & Kwon, incorporated by reference). A system and/or method for

biometrics-based facial feature extraction may be employed using a combination of disparity mapping, edge detection and filtering to determine co-ordinates for facial features in the region of interest (see, e.g., US patent 6,526,161 to Yan, incorporated by reference). A method for the automatic detection and tracking of personnel may utilize modules to track a users head or face (see, e.g., US patent 6,188,777, incorporated by reference). For example, a depth estimation module, a color segmentation module and/or a pattern classification module may be used. Data from each of these modules can be combined to assist in the identification of a user and the system can track and respond to a user's head or face in real-time.

The preferred and alternative embodiments may be applied in the field of digital photography. For example, automatic determination of main subjects in photographic images may be performed (see, e.g., US patent 6,282,317 to Luo et al., incorporated by reference). Regions of arbitrary shape and size may be extracted from a digital image. These may be grouped into larger segments corresponding to physically coherent objects. A probabilistic reasoning engine may then estimate the region which is most likely to be the main subject of the image.

Faces may be detected in complex visual scenes and/or in a neural network based face detection system, particularly for digital image processing in accordance with preferred or alternative embodiments herein (see, e.g., US patent 6,128,397 to Baluja & Rowley; and "Neural Network-Based Face Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 1, pages 23-28, January 1998 by the same authors, each reference being hereby incorporated by reference. In addition, an image may be rotated prior to the application of the neural network analysis in order to optimize the success rate of the neural-network based detection (see, e.g., US patent 6,128,397, incorporated by reference). This technique is particularly advantageous when faces are oriented vertically. Face detection in accordance with preferred and alternative embodiments, and which are particularly advantageous when a complex background is involved, may use one or more of skin color detection, spanning tree minimization and/or heuristic elimination of false positives (see, e.g., US patent 6,263,113 to Abdel-Mottaleb et al., incorporated by reference).

A broad range of techniques may be employed in image manipulation and/or image enhancement in accordance with preferred and alternative embodiments, may involve automatic, semi-automatic and/or manual operations, and are applicable to several fields of application. Some of the discussion that follows has been grouped into subcategories for ease of discussion, including (i) Contrast Normalization and Image Sharpening; (ii) Image Crop, Zoom and Rotate; (iii) Image Color Adjustment and Tone Scaling; (iv) Exposure Adjustment and Digital Fill Flash applied to a Digital Image; (v) Brightness Adjustment with Color Space Matching; and Auto-Gamma determination with Image Enhancement; (vi) Input/Output device characterizations to determine Automatic/Batch Image Enhancements; (vii) In-Camera Image Enhancement; and (viii) Face Based Image Enhancement. Other alternative embodiments may employ techniques provided at US application serial no. 10/608,784, filed June 26, 2003, which is hereby incorporated by reference.

Slide Show Based on One or More Image Features or Regions of Interest

Therefore in one embodiment, the creation of a slide show is based on the automated detection of face regions. In other embodiments, other image features, regions of interest (ROI) and/or characteristics are detected and employed in combination with detected face regions or independently to automatically construct a sophisticated slide show which highlights key features within a single image and or multiple images such as a sequence of images.

Examples of image features or regions, in addition to faces, are facial regions such as eyes, nose, mouth, teeth, cheeks, ears, eyebrows, forehead, hair, and parts or combinations thereof, as well as foreground and background regions of an image. Another example of a region of an image is a region that includes one or more faces and surrounding area of the image around the face or faces.

Separation of Foreground and Background Regions

Foreground and background regions may be advantageously separated in a preferred embodiment, which can include independent or separate detection, processing, tracking, storing, outputting, printing, cutting, pasting, copying, enhancing, upsampling, downsampling, fill flash processing, transforming, or other digital processing such as the exemplary processes provided in Tables I and II below. Independent transformations may be made to the foreground regions and the background regions. Such transformations are illustrated in the tables below. Table I lists several exemplary parameters that can be addressed regionally within an image or that can be addressed differently or adjusted different amounts at different regions within an image. With focus, selective out-of-focus regions can be created, while other regions are in focus. With saturation, selective reduction of color (gray scale) can be created, or different regions within an image can have different gray scales selected for them. With pixilation, selective reduction of amount of pixels per region can be applied. Sharpening can also be added region-by-region. With zooming, an image can be cropped to smaller regions of interest. With panning and tilting, it is possible to move horizontally and vertically, respectively, within an image. With dolly, foreshortening or a change of perspective is provided.

Table II illustrates initial and final states for different regions, e.g., foreground and background regions, within an image having processing applied differently to each of them. As shown, the initial states for each region are the same with regard to parameters such as focus, exposure sharpening and zoom, while addressing the regions differently during processing provides different final states for the regions. In one example, both the foreground and background regions are initially out of focus, while processing brings the foreground region into focus and leaves the background region out of focus. In another example, both regions are initially normal in focus, while processing takes the background out of focus and leaves the foreground in focus. In further examples, the regions are initially both normally exposed or both under exposed, and processing results in the foreground region being normally exposed and the background region being under exposed or over exposed. In another example, both regions are

initially normal sharpened, and processing results in over-sharpening of the foreground region and under-sharpening of the background region. In a further example, a full initial image with foreground and background is changed to a zoomed image to include only the foreground region or to include a cropped background region. In a further example, an initial image with normal background and foreground regions is changed to a new image with the foreground region zoomed in and the background region zoomed out.

Transformations can be reversed. For example, zoom-in or cropping may be reversed to begin with the cropped image and zoom out, or blurring that is sharpened may be reversed into an initial state of sharpening and final stages of blur, and so on with regard to the examples provided, or any permutations and any combinations of such transformations can be concatenated in various orders and forms (e.g., zoom and blur, blur and zoom).

TABLE I

Parameter	Effect
Focus	Create selective out-of-focus regions
Saturate	Create selective reduction of color (gray scale)
Pixelate	Selectively reduce amount of pixels per region
Sharpen	Add sharpening to regions
Zoom in	Crop image to smaller region of interest
Pan	Move horizontally across the image
Tilt	Move vertically up/down
Dolly	Change perspective, foreshortening

Examples include:

TABLE II

		Initial State		*Final State*		
		Foreground	**Background**	**Foreground**	**Background**	
		Out of Focus	Out of Focus	In Focus	Out of Focus	
		Normal In Focus	Normal In Focus	Normal In Focus	Out of Focus	
		Normal Good Exposure	Normal Good Exposure	Normal Good Exposure	Under Exposed	
		Under Exposed	Under Exposed	Normal Good Exposure	Under Exposed	
		Normal Good Exposure	Normal Good Exposure	Good Exposure	Over Exposed	
		Normal Sharpening	Normal Sharpening	Over sharpened	Under sharpened	
		Full Image with Foreground	Full Image with Background	Zoomed image to include only FG	Cropped Background	
		Normal	Normal	Zoomed in	Zoomed out (foreshortening)	

Alternatively, separated foreground/background regions may be further analyzed to determine their importance/relevance. In another embodiment, a significant background feature such as a sunset or a mountain may be incorporated as part of a slide show sequence. Foreground and background regions may be automatically separated, or semi-automatically, as described at U.S. patent application no. 11/217,788, Filed August 30, 2005, which is hereby incorporated by reference.

After separation of foreground and background regions it is also possible to calculate a depth map of the background regions. . By calculating such a depth map at the time that an image is acquired, it is possible to use additional depth map information to enhance the automatic generation of a slide show.

In the embodiment which preferably uses faces, yet is applicable to using other selected image features or regions, in case there are multiple faces detected, interesting "camera movement" can be simulated which includes panning/tilting from one face to another or zooming in-out onto a selection of faces.

While an exemplary drawings and specific embodiments of the present invention have been described and illustrated, it is to be understood that that the scope of the present invention is not to be limited to the particular embodiments discussed. Thus, the embodiments shall be regarded as illustrative rather than restrictive, and it should be understood that variations may be made in those embodiments by workers skilled in the arts without departing from the scope of the present invention as set forth in the claims that follow and their structural and functional equivalents.

In addition, in methods that may be performed according to the claims below and/or preferred embodiments herein, the operations have been described in selected typographical sequences. However, the sequences have been selected and so ordered for typographical convenience and are not intended to imply any particular order for performing the operations, unless a particular ordering is expressly provided or understood by those skilled in the art as being necessary.

What is claimed is:

1. A method of generating one or more new digital images using an original digitally-acquired image including a selected image feature, comprising:

(a) identifying one or more groups of pixels that correspond to a selected image feature within an original digitally-acquired image;

(b) selecting a portion of the original image that includes the one or more groups of pixels; and

(c) automatically generating values of pixels of one or more new images based on the selected portion in a manner which includes the selected image feature within the one or more new images.

2. The method of claim 1, further comprising gradually displaying a transformation between said original digitally-acquired image and said one or more new images.

3. The method claim 2, further comprising adjusting parameters of said transformation between said original digitally-acquired image and one or more new images.

4. The method of claim 3, wherein said parameters of said transformation between said original digitally-acquired image and one or more new images include timing or blending or a combination thereof.

5. The method claim 4, wherein said blending includes dissolving, flying, swirling, appearing, flashing, or screening, or combinations thereof.

6. The method of claim 2, wherein said transformation is different between the selected portion and remaining portions of the image.

7. The method of claim 6, further comprising:

(d) determining one or more further new images each including a new group of pixels corresponding to the selected image feature; and

(e) automatically creating a simulated camera movement.

8. The method of claim 7, wherein said simulated camera movement includes panning, tilting, or perspective modification, or combinations thereof.

9. The method of claim 1, wherein said selected image feature comprises one or more faces.

10. The method of claim 1, wherein said selected image feature comprises a human subject.

11. The method of claim 1, wherein said selected image feature comprises an animal.

12. The method of claim 11, wherein said animal comprises a domesticated pet.

13. The method of claim 1, wherein the selected image feature comprises a foreground region or a background region.

14. The method of claim 13, wherein the foreground region is determined by detection of a face.

15. The method of claim 13, wherein the foreground region is determined by defining said selected image feature within an original image to be local sharpness, relative exposure, local color clustering, or local saturation, or combinations thereof.

16. The method of claim 13, wherein the determining of the foreground region by defining said selected image feature within an original image comprises determining a depth of focus.

17. The method of claim 13, further comprising visually separating the foreground region and the background region within the one or more new images.

18. The method of claim 17, further comprising calculating a depth map of the background region.

19. The method of claim 13, further comprising independently processing the foreground region or the background region, or both.

20. The method of claim 19, wherein at least one of said new images comprises an independently processed background region or foreground region or both.

21. The method of claim 20, wherein the independent processing comprises focusing, saturating, pixilating, sharpening, zooming, panning, tilting, cropping, geometrically distorting, or exposing, or combinations thereof.

22. The method of claim 16, further comprising determining a relevance or importance, or both, of the foreground region or the background region, or both.

23. The method of claim 1, wherein the identifying comprises identifying one or more groups of pixels that correspond to two or more selected image features within the original digitally-acquired image; and wherein the automatic generating is in a manner which includes at least one of the two or more selected image features within the one or more new images or a panning intermediate image between two of the selected image features, or a combination thereof.

24. The method of claim 23, further comprising visually shifting a center of interest between the two or more identified image features using simulated camera movement.

25. The method of claim 23, wherein said simulated camera movement comprises identifying a direction parameter between said two of the identified image features.

26. The method of claim 25, wherein said identifying camera movement is determined based on the determination of a vanishing point of the image.

27. The method of claim 25, wherein said identifying camera movement is determined based on spatial orientation and line of sight of one or more faces in said image.

28. A method as recited in claim 1, wherein said identifying one or more groups of pixels is performed within a digital acquisition device.

29. A method as recited in claim 28, wherein said identifying one or more groups of pixels within a digital acquisition device is performed based on information from one or more preview images.

30. A method as recited in claim 28, wherein said identifying one or more groups of pixels within a digital acquisition device is performed using nonimage data.

31. A method as recited in claim 30, wherein said non-image data comprises one or more acquisition parameters.

32. A method as recited in claim 28, wherein said generating values of pixels of one or more new images is performed within a digital acquisition device.

33. A method as recited in claim 28, wherein said generating values of pixels of one or more new images is performed by an external device to said digital acquisition device.

34. A method as recited in claim 1, further comprising:

(d) saving a transformation between said original digitally-acquired image and said one or more new images as a movie clip.

35. A method of generating one or more new digital images using an original digitally-acquired image including a background region or a foreground region, or both, comprising:

(a) identifying one or more groups of pixels that correspond to a background region or a foreground region, or both, within an original digitally-acquired image;

(b) selecting a portion of the original image that includes the one or more groups of pixels; and

(c) automatically generating values of pixels of one or more new images based on the selected portion in a manner which includes the background region or the foreground region, or both, within each of the one or more new images.

36. The method of claim 35, further comprising separating the foreground region and the background region within the original image or the one or more new images or combinations thereof.

37. The method of claim 36, further comprising calculating a depth map of the background region or the foreground region or both.

38. The method of claim 35, further comprising independently processing the foreground region or the background region, or both.

39. The method of claim 38, wherein at least one of said new images comprises an independently processed background region or foreground region or both.

40. The method of claim 39, wherein the independent processing comprises focusing, saturating, pixilating, sharpening, zooming, panning, tilting, geometrical distorting, cropping, exposing or combinations thereof.

41. The method of claim 35, further comprising determining a relevance or importance, or both, of the foreground region or the background region, or both.

42. One or more processor readable storage devices having processor readable code embodied thereon, said processor readable code for programming one or more processors to perform a method of generating one or more new digital images using an original digitally-acquired image including a selected image feature, the method comprising:

(a) identifying one or more groups of pixels that correspond to a selected image feature within an original digitally-acquired image;

(b) selecting a portion of the original image that includes the one or more groups of pixels; and

(c) automatically generating values of pixels of one or more new images based on the selected portion in a manner which includes the selected image feature within the one or more new images.

43. One or more storage devices as recited in claim 42, the method further comprising gradually displaying a transformation between said original digitally-acquired image and said one or more new images.

44. One or more storage devices as recited in claim 43, the method further comprising adjusting parameters of said transformation between said original digitally-acquired image and one or more new images.

45. One or more storage devices as recited in claim 44, wherein said parameters of said transformation between said original digitally-acquired image and one or more new images include timing or blending or a combination thereof.

46. One or more storage devices as recited in claim 45, wherein said blending includes dissolving, flying, swirling, appearing, flashing, or screening, or combinations thereof.

47. One or more storage devices of claim 43, wherein said transformation is different between the selected portion and remaining portions of the image.

48. One or more storage devices of claim 47, the method further comprising:

(d) determining one or more further new images each including a new group of pixels corresponding to the selected image feature; and

(e) automatically creating a simulated camera movement.

49. One or more storage devices of claim 48, wherein said simulated camera movement includes panning, tilting, or perspective modification, or a combinations thereof.

50. One or more storage devices of claim 42, wherein said selected image feature comprises one or more faces.

51. One or more storage devices of claim 42, wherein said selected image feature comprises a human subject.

52. One or more storage devices of claim 42, wherein said selected image feature comprises an animal.

53. One or more storage devices of claim 52, wherein said animal comprises a domesticated pet.

54. One or more storage devices of claim 42, wherein the selected image feature comprises a foreground region or a background region.

55. One or more storage devices of claim 54, wherein the foreground region is determined by detection of a face.

56. One or more storage devices of claim 54, wherein the foreground region is determined by defining said selected image feature within an original image to be local sharpness, relative exposure, local color clustering, or local saturation, or combinations thereof.

57. One or more storage devices of claim 54, wherein the determining of the foreground region by defining said selected image feature within an original image comprises determining a depth of focus.

58. One or more storage devices of claim 54, the method further comprising visually separating the foreground region and the background region within the one or more new images.

59. One or more storage devices of claim 58, the method further comprising calculating a depth map of the background region.

60. One or more storage devices of claim 54, the method further comprising independently processing the foreground region or the background region, or both.

61. One or more storage devices of claim 60, wherein at least one of said new images comprises an independently processed background region or foreground region or both.

62. One or more storage devices of claim 61, wherein the independent processing comprises focusing, saturating, pixilating, sharpening, zooming, panning, tilting, cropping, geometrically distorting, or exposing, or combinations thereof.

63. One or more storage devices of claim 57, the method further comprising determining a relevance or importance, or both, of the foreground region or the background region, or both.

64. One or more storage devices of claim 42, wherein the identifying comprises identifying one or more groups of pixels that correspond to two or more selected image features within the original digitally-acquired image; and wherein the automatic generating is in a manner which includes at least one of the two or more selected image features within the one or more new images or a panning intermediate image between two of the selected image features, or a combination thereof.

65. One or more storage devices of claim 64, the method further comprising visually shifting a center of interest between the two or more identified image features using simulated camera movement. .

66. One or more storage devices of claim 64, wherein said simulated camera movement comprises identifying a direction parameter between said two of the identified image features.

67. One or more storage devices of claim 66, wherein said identifying camera movement is determined based on the determination of a vanishing point of the image.

68. One or more storage devices of claim 66, wherein said identifying camera movement is determined based on spatial orientation and line of sight of one or more faces in said image.

69. One or more storage devices as recited in claim 42, wherein said identifying one or more groups of pixels is performed within a digital acquisition device.

70. One or more storage devices as recited in claim 69, wherein said identifying one or more groups of pixels within a digital acquisition device is performed based on information from one or more preview images.

71. One or more storage devices as recited in claim 69, wherein said identifying one or more groups of pixels within a digital acquisition device is performed using non-image data.

72. One or more storage devices as recited in claim 71, wherein said non-image data comprises one or more acquisition parameters.

73. One or more storage devices as recited in claim 69, wherein said generating values of pixels of one or more new images is performed within a digital acquisition device.

74. One or more storage devices as recited in claim 69, wherein said generating values of pixels of one or more new images is performed by an external device to said digital acquisition device.

75. One or more storage devices as recited in claim 42, the method further comprising:

(d) saving a transformation between said original digitally-acquired image and said one or more new images as a movie clip.

76. One or more processor readable storage devices having processor readable code embodied thereon, said processor readable code for programming one or more processors to perform a method of generating one or more new digital images using an original digitally-acquired image including a background region or a foreground region, or both, the method comprising:

(a) identifying one or more groups of pixels that correspond to a background region or a foreground region, or both, within an original digitally-acquired image;

(b) selecting a portion of the original image that includes the one or more groups of pixels; and

(c) automatically generating values of pixels of one or more new images based on the selected portion in a manner which includes the background region or the foreground region, or both, within each of the one or more new images.

77. One or more storage devices of claim 76, the method further comprising separating the foreground region and the background region within the original image or the one or more new images or combinations thereof.

78. One or more storage devices of claim 77, the method further comprising calculating a depth map of the background region or the foreground region or both.

79. One or more storage devices of claim 76, the method further comprising independently processing the foreground region or the background region, or both.

80. One or more storage devices of claim 79, wherein at least one of said new images comprises an independently processed background region or foreground region or both.

81. One or more storage devices of claim 80, wherein the independent processing comprises focusing, saturating, pixilating, sharpening, zooming, panning, tilting, geometrical distorting, cropping, exposing or combinations thereof.

82. One or more storage devices of claim 76, the method further comprising determining a relevance or importance, or both, of the foreground region or the background region, or both.

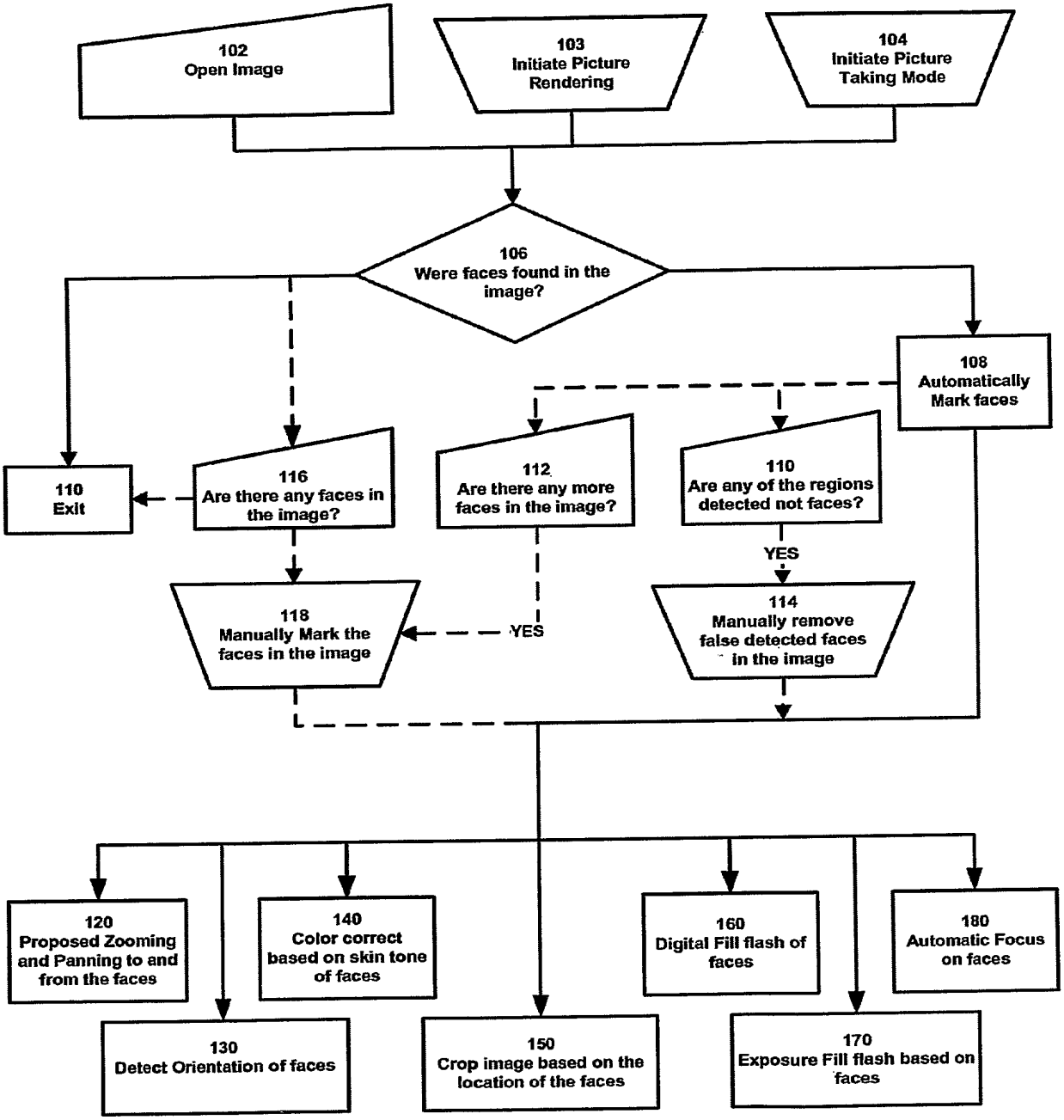

102
Open Image
103
Initiate Picture Rendering
104
Initiate Picture Taking Mode
106
Were faces found in the image?
108
Automatically Mark faces
110
Exit
116
Are there any faces in the image?
112
Are there any more faces in the image?
110
Are any of the regions detected not faces?
YES
118
Manually Mark the faces in the image
YES
114
Manually remove false detected faces in the image
120
Proposed Zooming and Panning to and from the faces
140
Color correct based on skin tone of faces
160
Digital Fill flash of faces
180
Automatic Focus on faces
130
Detect Orientation of faces
150
Crop image based on the location of the faces
170
Exposure Fill flash based on faces

Figure 1a

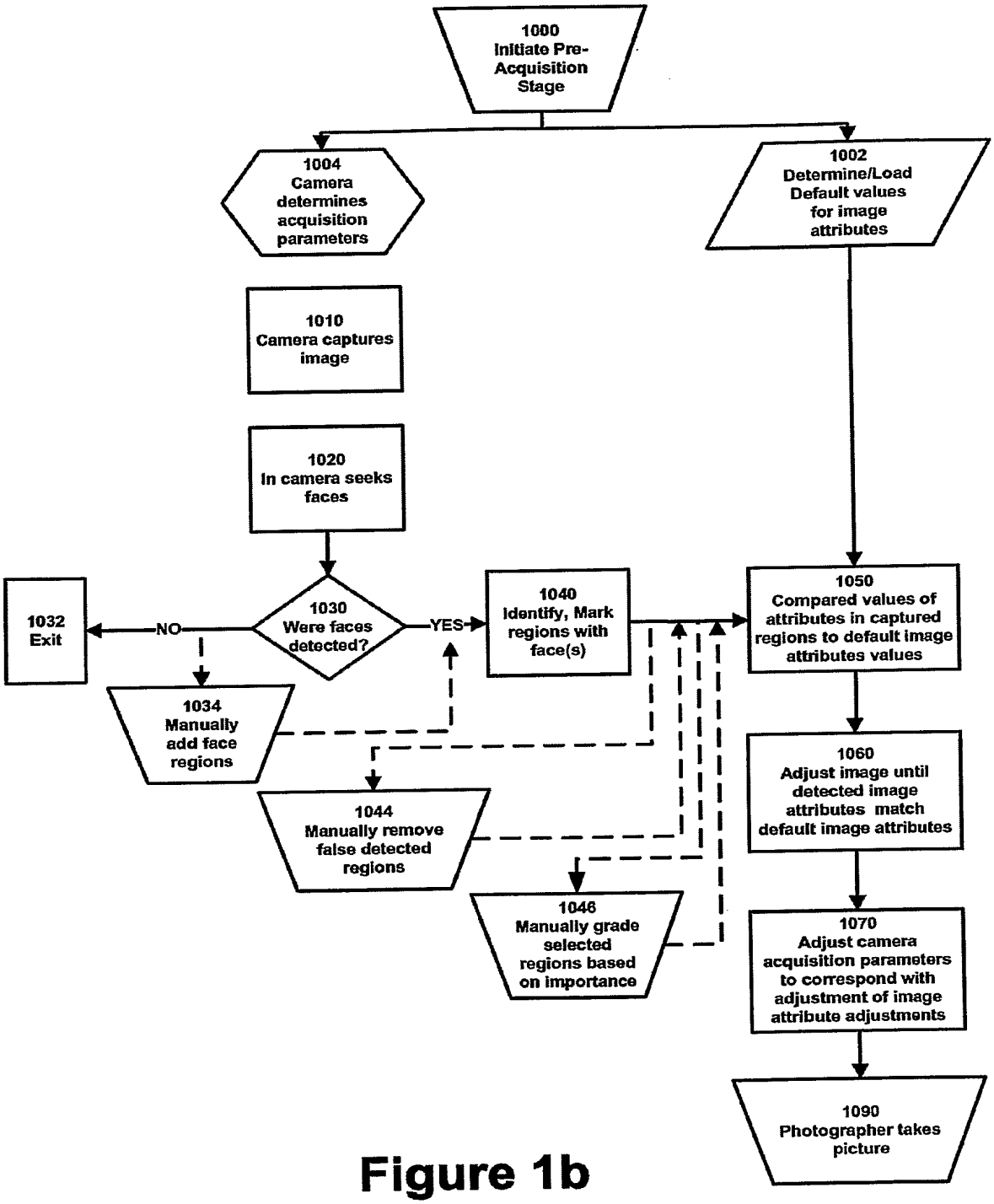

1000
Initiate Pre-Acquisition Stage
1004
Camera determines acquisition parameters
1002
Determine/Load Default values for image attributes
1010
Camera captures image
1020
In camera seeks faces
1030
Were faces detected?
NO
YES
1032
Exit
1034
Manually add face regions
1040
Identify, Mark regions with face(s)
1044
Manually remove false detected regions
1046
Manually grade selected regions based on importance
1050
Compared values of attributes in captured regions to default image attributes values
1060
Adjust image until detected image attributes match default image attributes
1070
Adjust camera acquisition parameters to correspond with adjustment of image attribute adjustments
1090
Photographer takes picture

Figure 1b

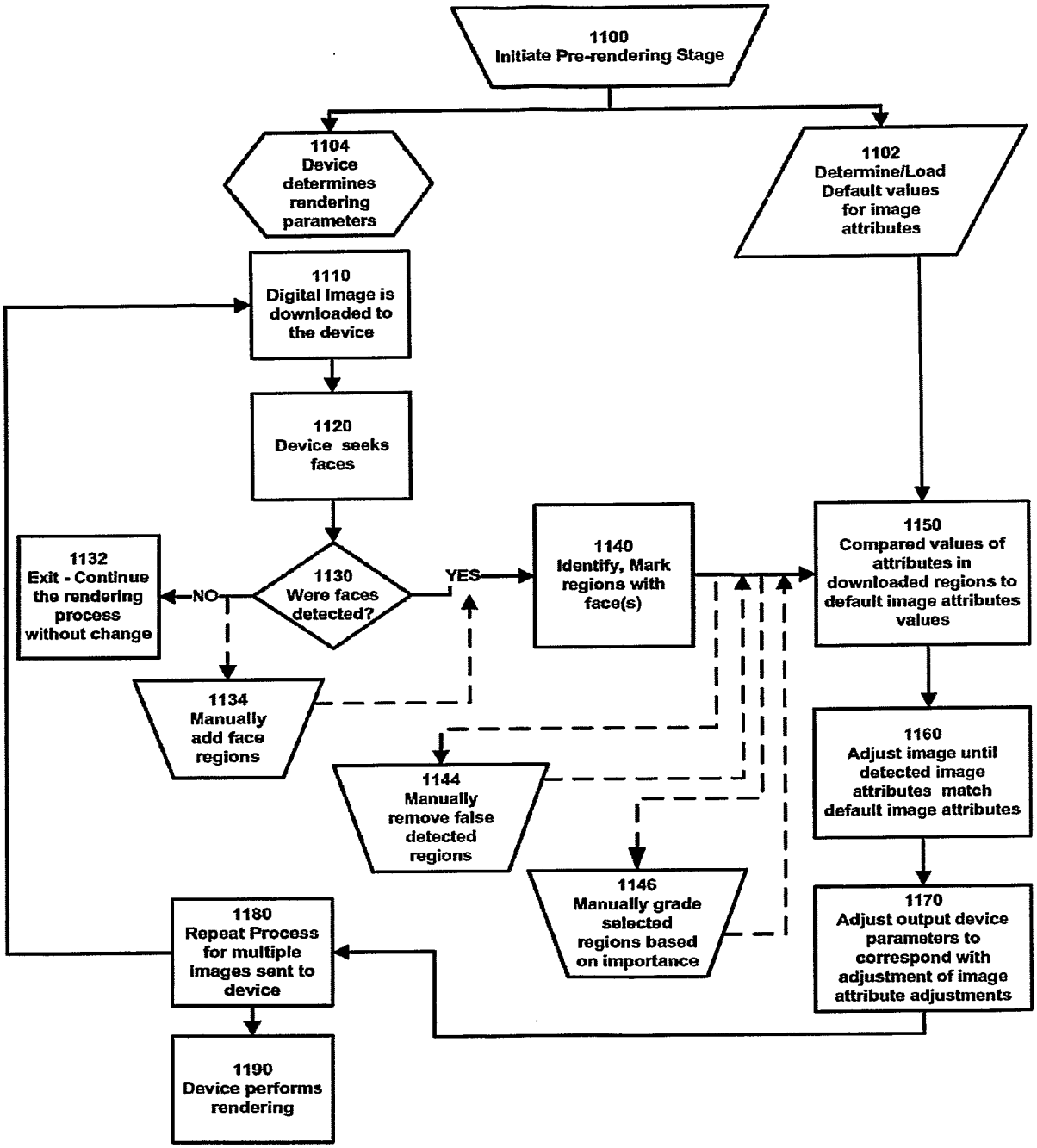

1100
Initiate Pre-rendering Stage
1104
Device determines rendering parameters
1102
Determine/Load Default values for image attributes
1110
Digital Image is downloaded to the device
1120
Device seeks faces
1130
Were faces detected?
YES
NO
1132
Exit - Continue the rendering process without change
1134
Manually add face regions
1140
Identify, Mark regions with face(s)
1144
Manually remove false detected regions
1146
Manually grade selected regions based on importance
1150
Compared values of attributes in downloaded regions to default image attributes values
1160
Adjust image until detected image attributes match default image attributes
1170
Adjust output device parameters to correspond with adjustment of image attribute adjustments
1180
Repeat Process for multiple images sent to device
1190
Device performs rendering

Figure 1c

Figure 2a

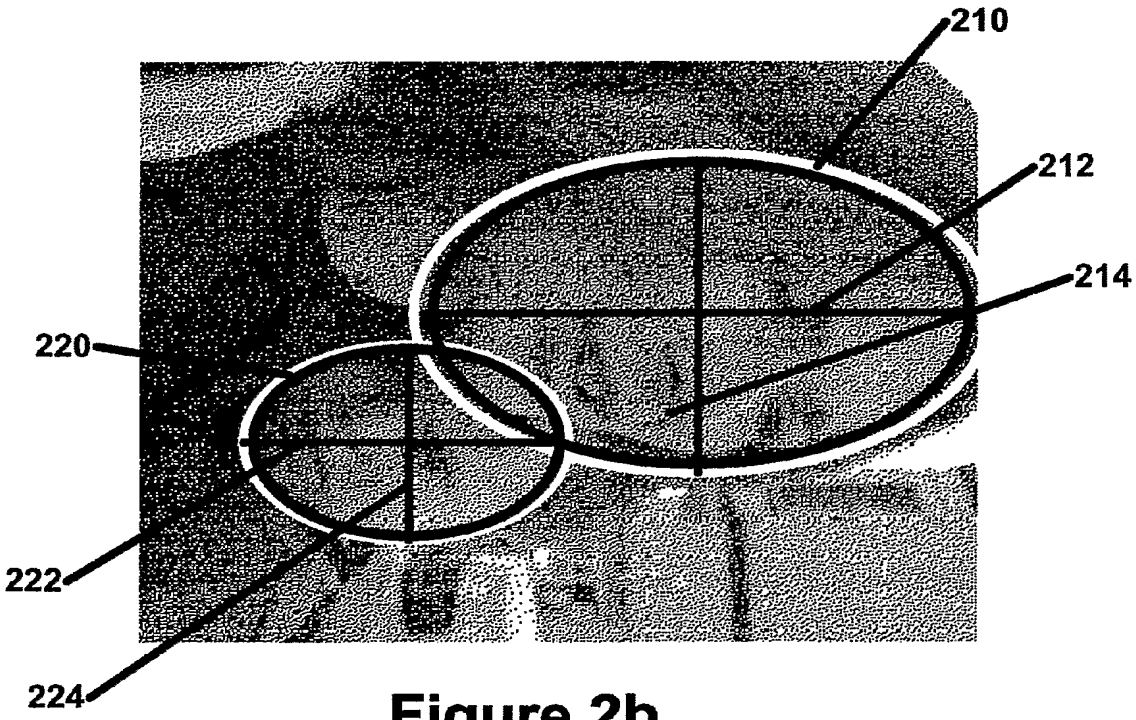


Figure 2b

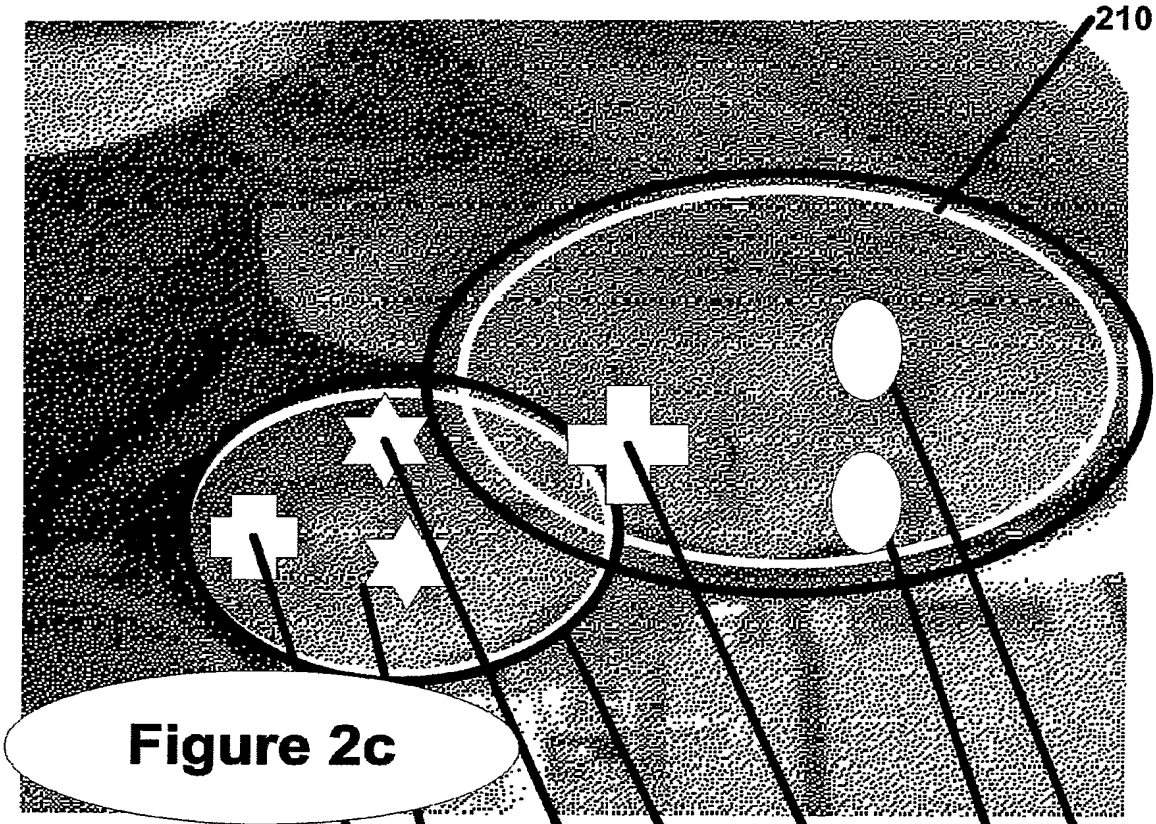


Figure 2c

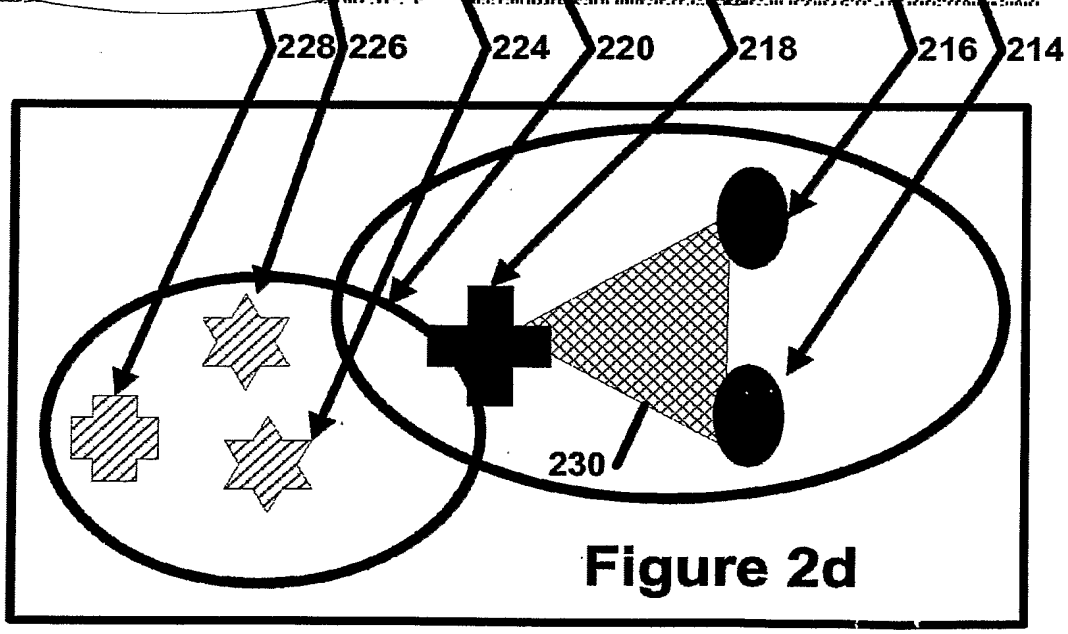


Figure 2d

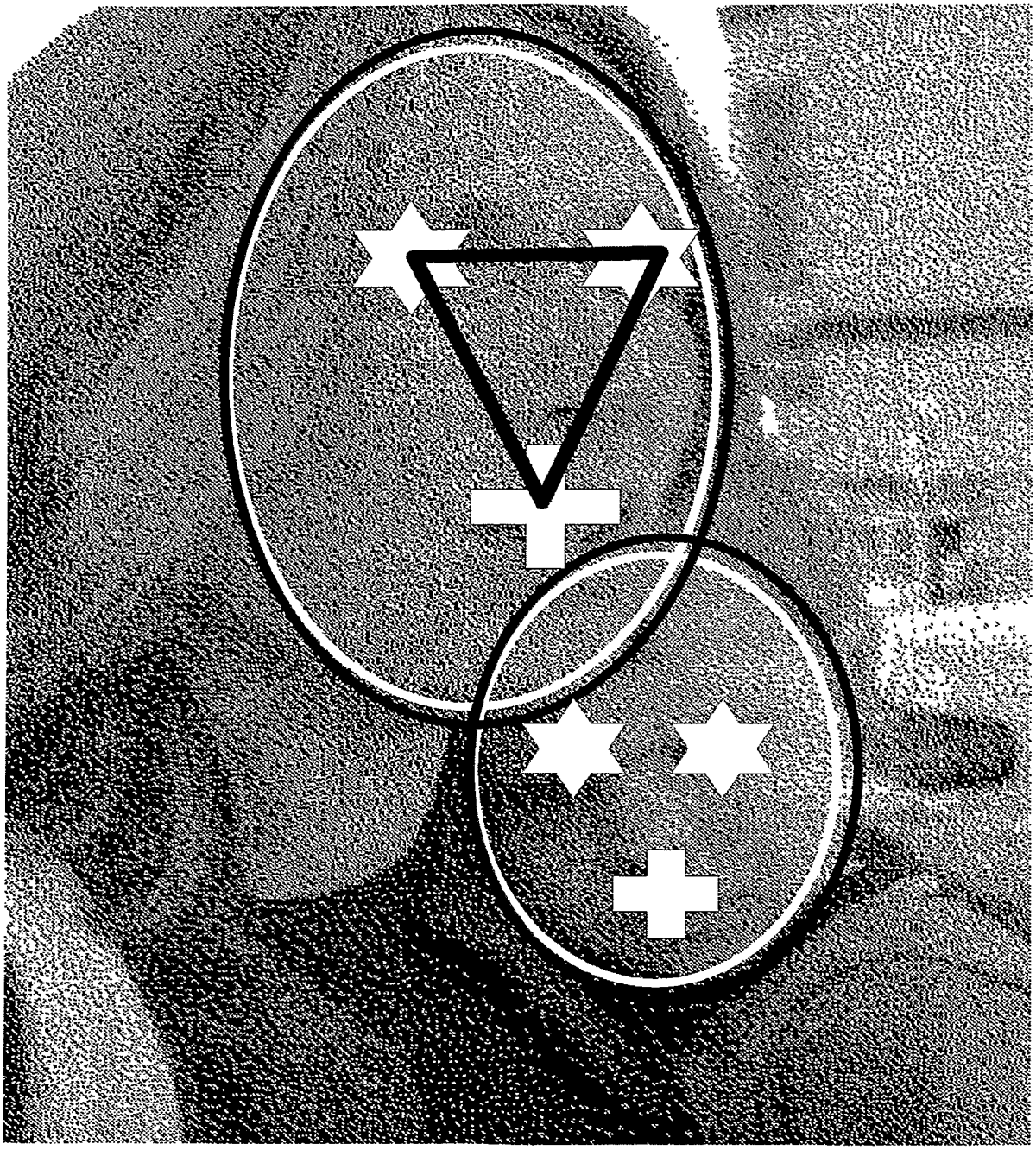
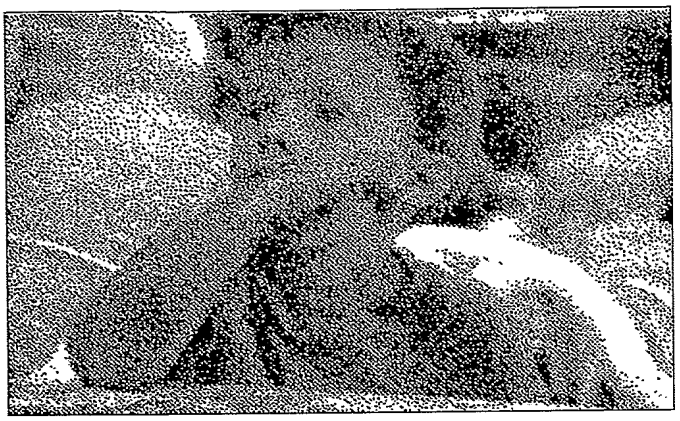


Figure 2e

Figure 3a



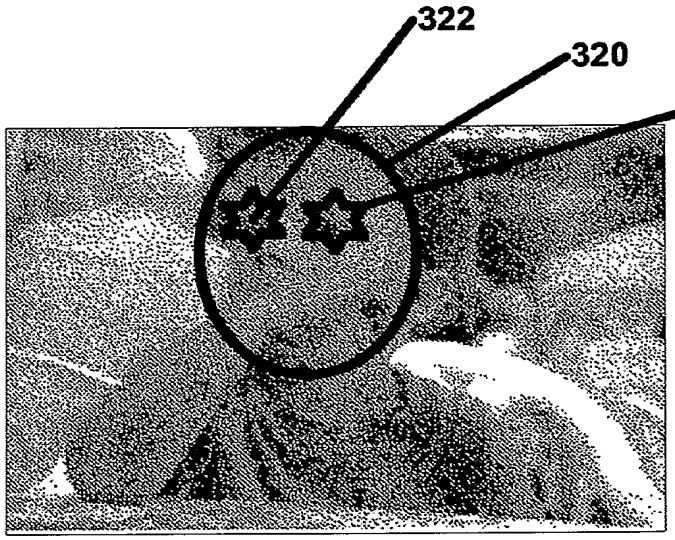


Figure 3b

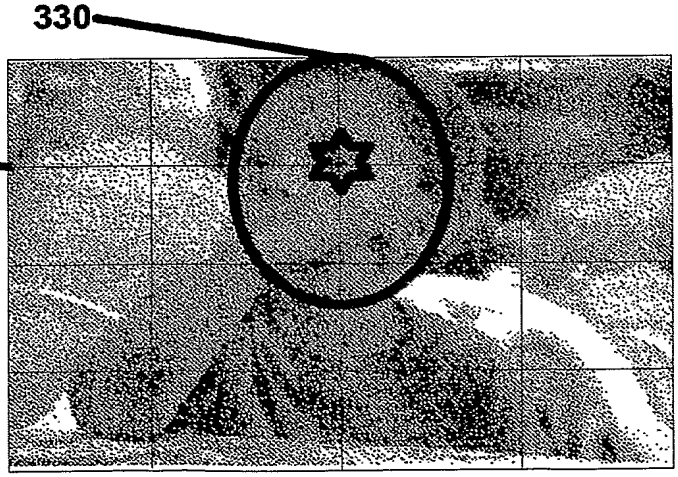


Figure 3c

Figure 3d

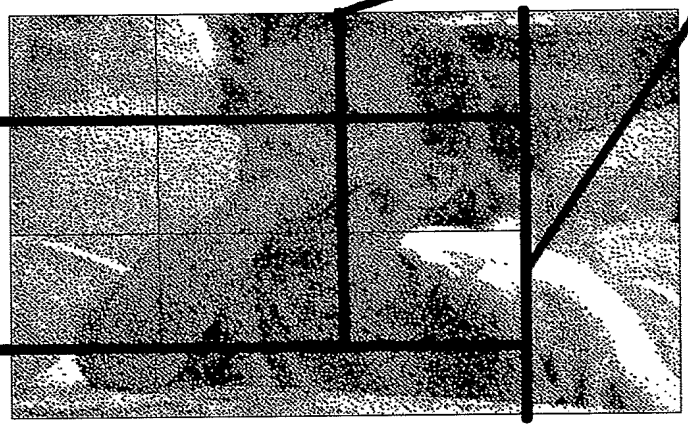


Figure 3e

Figure 3f

Figure 4a

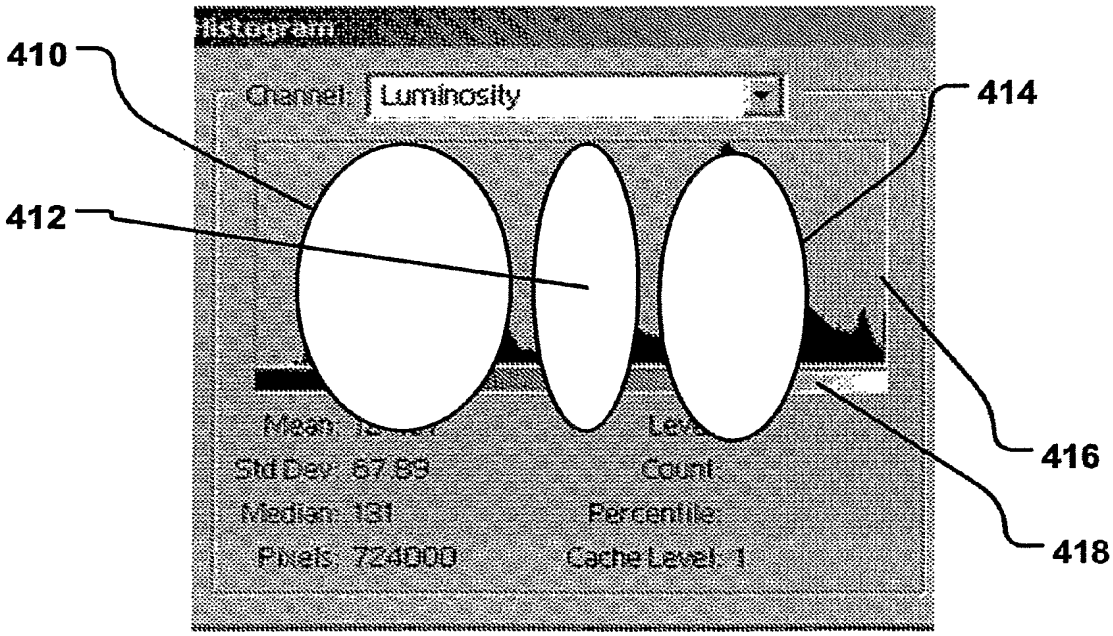


Figure 4b

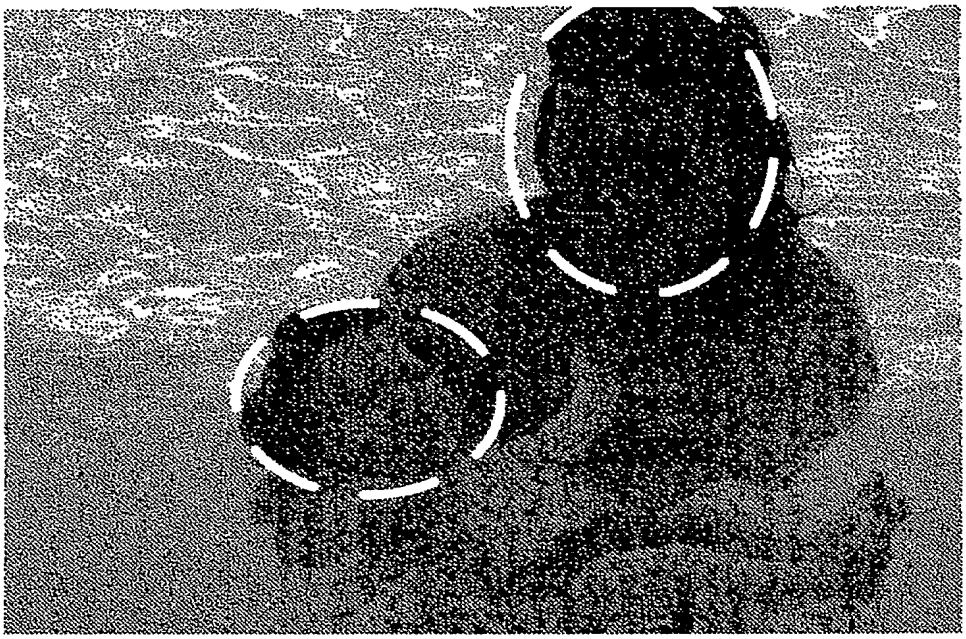


Figure 4c

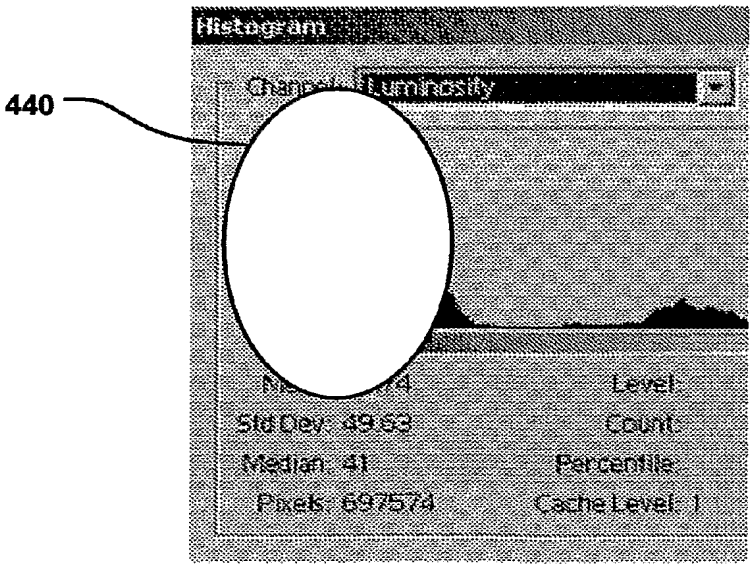
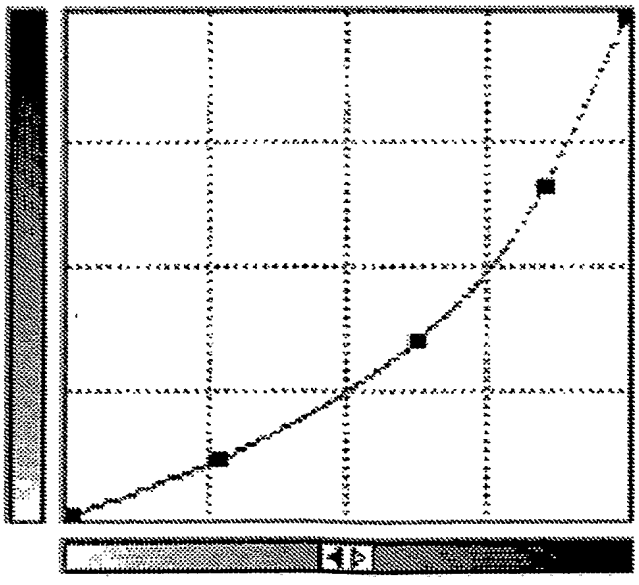


Figure 4d

Figure 4e

Figure 4f



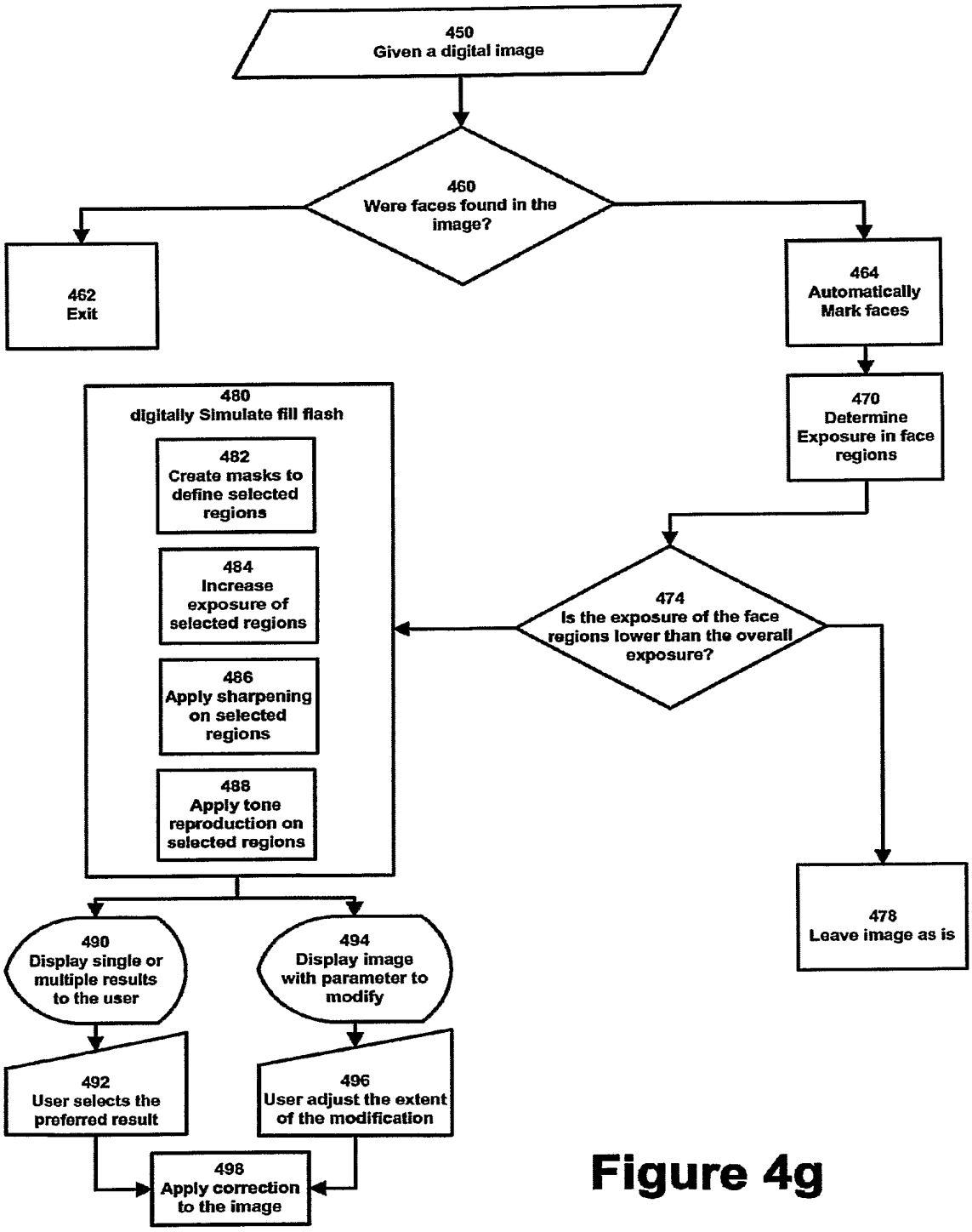

450
Given a digital image
460
Were faces found in the image?
462
Exit
464
Automatically Mark faces
470
Determine Exposure in face regions
474
Is the exposure of the face regions lower than the overall exposure?
478
Leave image as is
480
digitally Simulate fill flash
482
Create masks to define selected regions
484
Increase exposure of selected regions
486
Apply sharpening on selected regions
488
Apply tone reproduction on selected regions
490
Display single or multiple results to the user
492
User selects the preferred result
494
Display image with parameter to modify
496
User adjust the extent of the modification
498
Apply correction to the image

Figure 4g

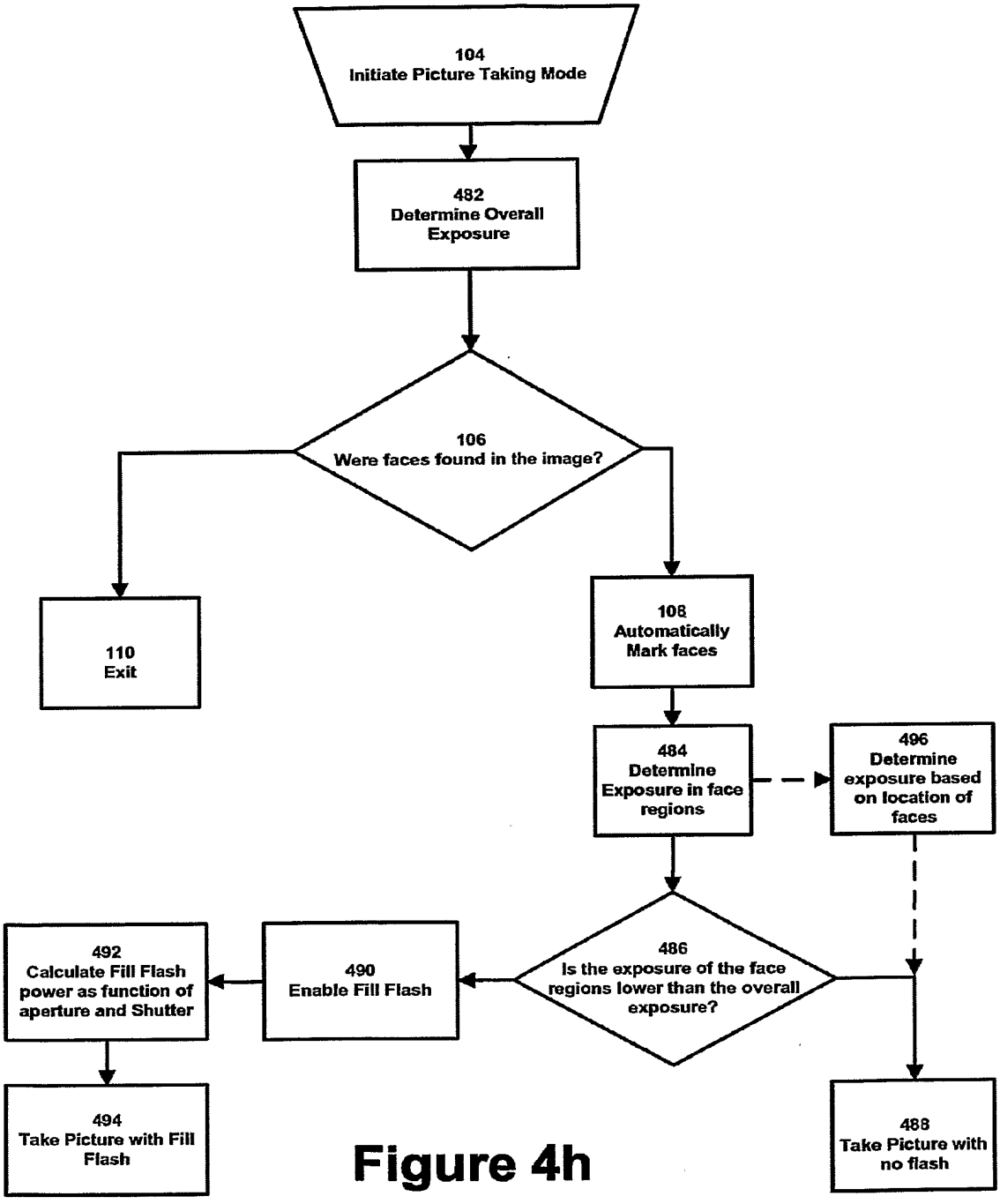

104
Initiate Picture Taking Mode
482
Determine Overall Exposure
106
Were faces found in the image?
110
Exit
108
Automatically Mark faces
484
Determine Exposure in face regions
496
Determine exposure based on location of faces
486
Is the exposure of the face regions lower than the overall exposure?
490
Enable Fill Flash
492
Calculate Fill Flash power as function of aperture and Shutter
494
Take Picture with Fill Flash
488
Take Picture with no flash

Figure 4h

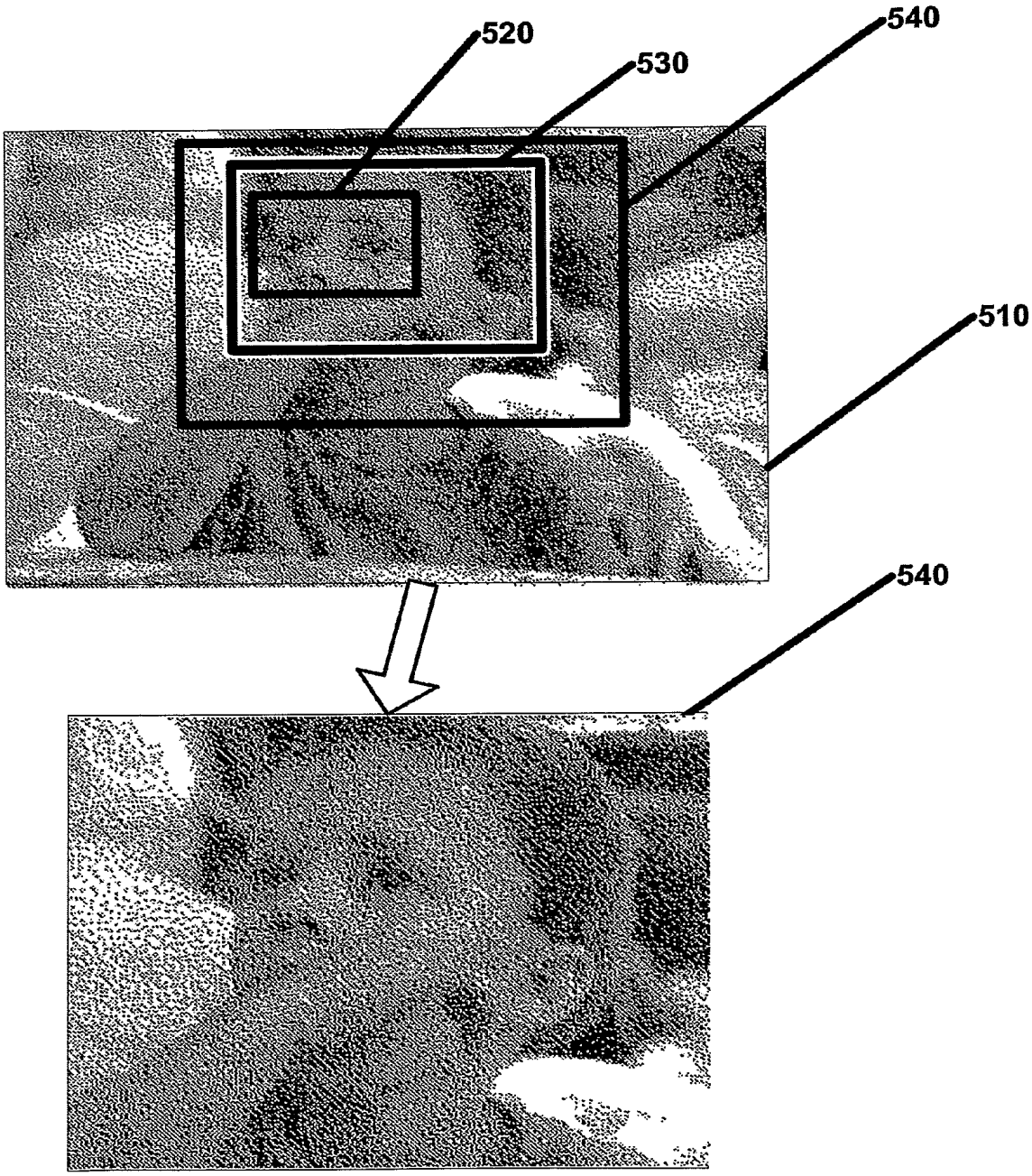


Figure 5

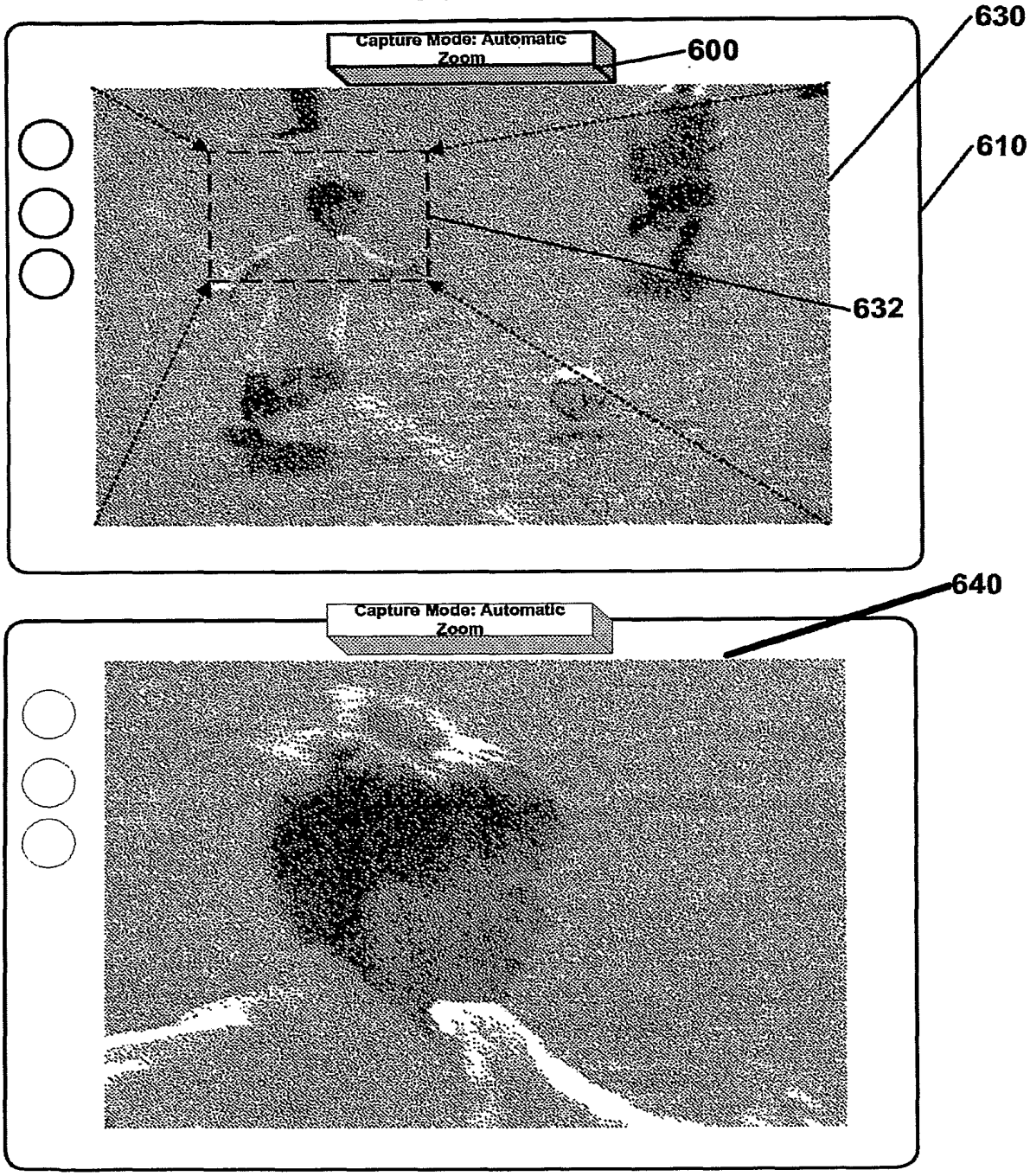

Capture Mode: Automatic Zoom
600
630
610
632
Capture Mode: Automatic Zoom
640

Figure 6

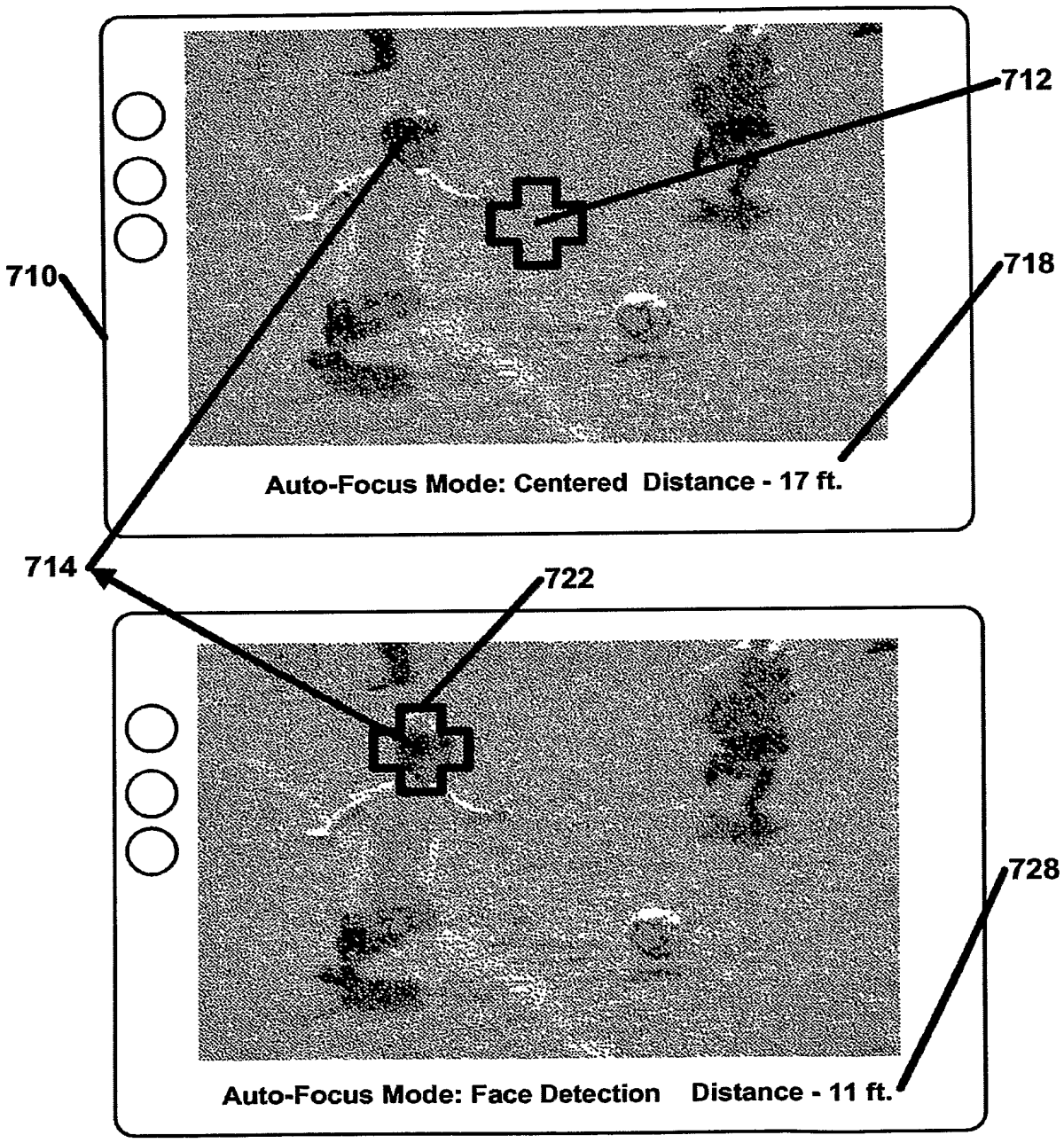

712
710
718
Auto-Focus Mode: Centered Distance - 17 ft.
714
722
728
Auto-Focus Mode: Face Detection Distance - 11 ft.

Figure 7a

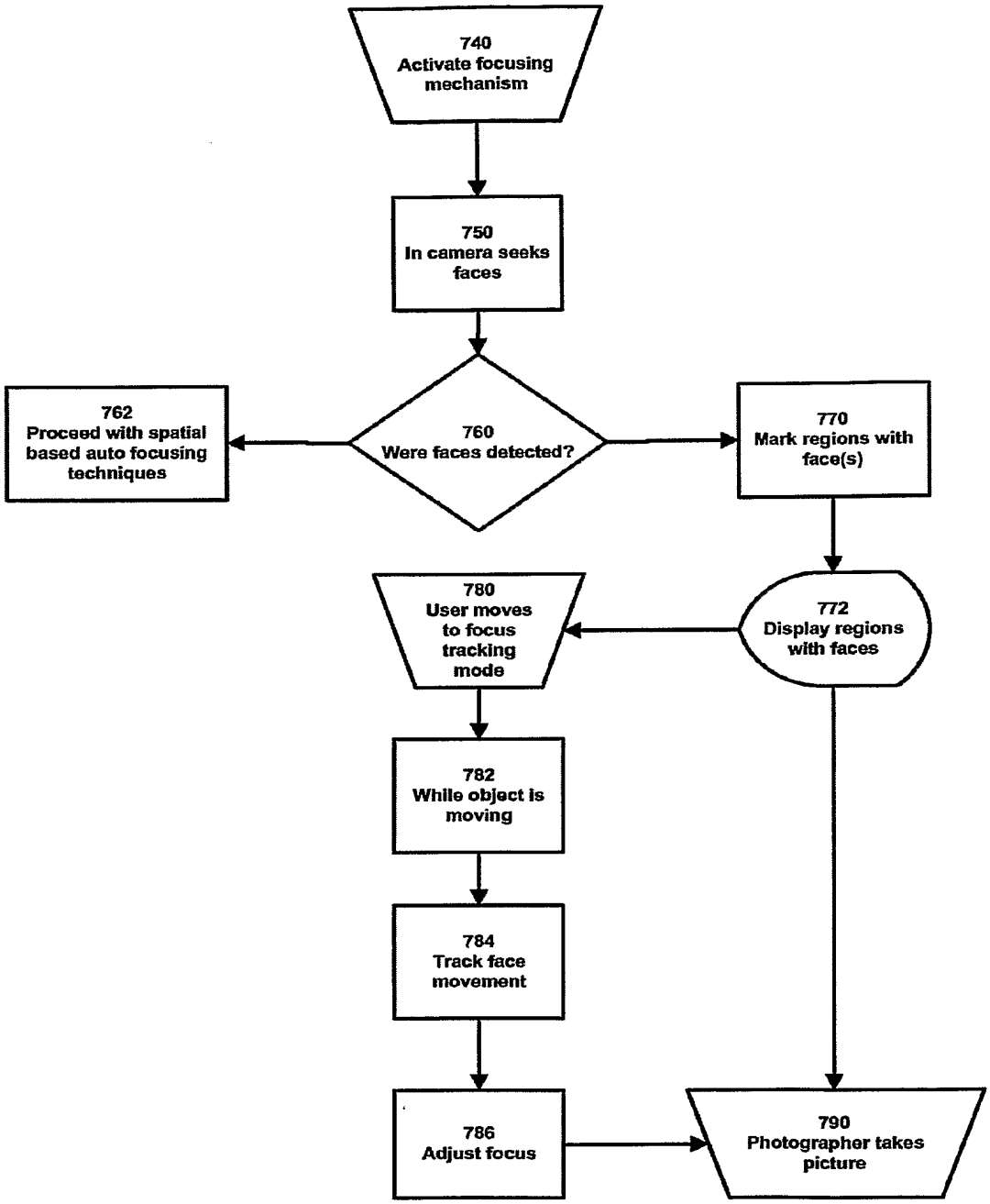

740
Activate focusing mechanism
750
In camera seeks faces
760
Were faces detected?
762
Proceed with spatial based auto focusing techniques
770
Mark regions with face(s)
772
Display regions with faces
780
User moves to focus tracking mode
782
While object is moving
784
Track face movement
786
Adjust focus
790
Photographer takes picture

Figure 7b

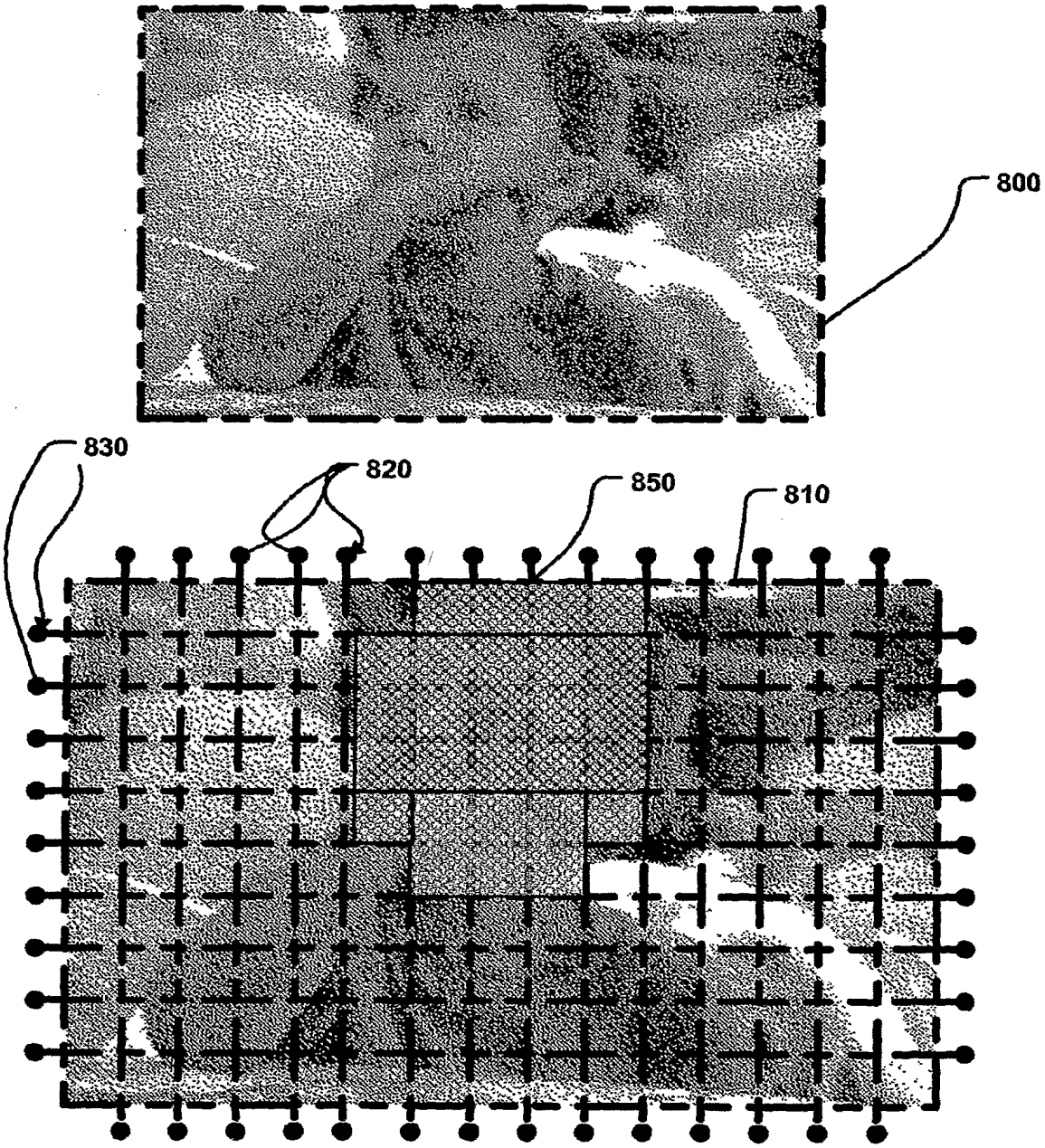


Figure 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US06/21393

A. CLASSIFICATION OF SUBJECT MATTER
IPC: **G06K 9/00**(2006.01);**G06K 9/34**(2006.01);**G06K 9/36**(2006.01);**G06K 9/32**(2006.01);**G06K 9/40**(2006.01)

USPC: 382/103,173,254,282,296
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 382/103,173,254,282,296

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 20030142209 (Yamazaki et al) 31 July 2003 (31.07.2003), fig. 3	24-27, 34, 65- 68 and 75
Y	US 20030123751(Krishnamurthy et al) 03 July 2003 (03.07.2003), entire publication	1-80
Y	US 4970663 (Bedell et al) 13 November 1990 (13.11.1990), col. 1, line 39 and col 4, lines 19-20	1-80
Y	US 7110569 (Brodsky et al) 19 September 2006 (19.10.2006), col. 6, line 5	12,53
Y	US 5384615 (Hsieh et al) 24 January 1995 (24.01.1995), col. 4, lines 48-49	16,22,57,63

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
24 January 2007 (24.01.2007)

Date of mailing of the international search report
29 MAR 2007

Name and mailing address of the ISA/US
Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
Facsimile No. (571) 273-3201

Authorized officer
Joseph Mancuso
Jacqueline A. Whitfield
Special Project Asst.
Telephone No. (571)272-8535

Form PCT/ISA/210 (second sheet) (April 2005)